

**Funktionelle Analyse kleiner, nichtkodierender RNAs in den  
Organellen von *Plasmodium falciparum* und Charakterisierung  
neuer RNA-Bindeproteine in Apicomplexa**

**DISSERTATION**

zur Erlangung des akademischen Grades

**D o c t o r r e r u m n a t u r a l i u m**

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der Lebenswissenschaftlichen Fakultät  
der Humboldt-Universität zu Berlin

von

Master of Science, Arne Thomas Hillebrand,

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Bernhard Grimm

Gutachter/innen: 1. Prof. Dr. Christian Schmitz-Linneweber

2. Prof. Dr. Kai Matuschewski

3. Prof. Dr. Volker Knoop

Tag der mündlichen Prüfung: 11.07.2019



**Für meine Familie, Diesseits und Jenseits**





## Zusammenfassung

Die Infektionskrankheit Malaria wird von einzelligen Parasiten der Gattung *Plasmodium* verursacht und stellt vor allem im südlichen Afrika eine große Herausforderung für das Gesundheitssystem dar. Die Zellen der Parasiten enthalten zwei endosymbiontische Organellen, den Apicoplast und das Mitochondrium. Beide Organellen besitzen ein reduziertes Genom. Die Struktur des mitochondrialen Genoms ist ungewöhnlich. Mit nur 6 kb gehört es zu den kleinsten beschriebenen Genomen und enthält neben drei proteinkodierende Gene auch 34 kleine rRNA-Gene. Um das Genom zu exprimieren wird eine Vielzahl von kernkodierten Faktoren benötigt. Die Regulation der Expression, die Prozessierung der polycistronischen Primärtranskripte und die Regulation des RNA-Metabolismus des Mitochondriums ist jedoch weitestgehend unbekannt.

In dieser Arbeit konnten kurze RNAs in den Mitochondrien von *P. falciparum* mittels Hochdurchsatzsequenzierung identifiziert werden. Ihre stabile Akkumulation wurde mittels RNA-Gelblothybridisierung bestätigt. Ein Teil dieser kurzen RNAs konnte mRNAs zugeordnet werden, während ein anderer Teil an den 5'-Enden von ribosomalen Fragmenten akkumuliert. Solche RNA-Akkumulationen an Transkriptenden werden in den Organellen höherer Pflanzen durch PPR-Proteine (*Pentatricopeptide repeat*) verursacht und als clustered organellar short RNAs (cosRNA) bezeichnet.

Um zu untersuchen, ob in *P. falciparum* PPR-ähnliche, helikale Proteine vorhanden sind, wurde genomweit nach Proteinen mit repetitiven, helikalen Elementen gesucht. Dabei konnte eine vorher unbekannte Proteinfamilie identifiziert werden, die aufgrund ihrer 37 Aminosäure langen Motive und der strukturellen Ähnlichkeit zu den PPR-Motiven *Heptatricopeptide repeat* Proteine (HPR) genannt wurde. Die Lokalisationsvorhersage der 22 HPR-Proteine in *P. falciparum* lässt eine überwiegend mitochondriale Lokalisation vermuten. In *P. berghei*, dem Malariaerreger in Nagetieren, konnte für 7 HPR-Proteine die mitochondriale Lokalisation betätigt werden. Außerdem zeigten Deletionsversuche, dass die meisten HPR-Proteine in den Blutstadien essentiell sind. In *in vitro* RNA-Bindestudien konnte für ein rekombinantes HPR-Protein eine unspezifische

Interaktion mit mitochondrialen Transkripten nachgewiesen werden, während keine Bindung an DNA erfolgt. Eine breite Suche in verschiedenen phylogenetischen Gruppen zeigte, dass HPR-Proteine in verschiedensten eukaryotischen Taxa vorhanden sind, mithin früh in der Evolution der eukaryotischen Zelle entstanden sind. Im Menschen konnten 6 Proteine mit HPR-Motiven identifiziert werden. Diese gehören alle zur FASTK-Proteinfamilie, einer mitochondrial lokalisierten Proteinfamilie, die am RNA-Metabolismus des menschlichen Mitochondriums beteiligt ist.

## Summary

Malaria is caused by a single celled parasite of the genus *Plasmodium*. Especially in Sub-Saharan Africa, -this disease is a huge challenge for the health system. The cells of the parasites contain two organelles of endosymbiotic origin, the apicoplast and the mitochondrion. Both organelles still contain a reduced genome. For the expression of the genome, the organelles depends on a large set of nuclear encoded proteins. The mitochondrial genome has a unique structure. With only 6 kb it is one of the smallest genomes discovered to date and it contains only three protein coding genes along with 34 small ribosomal genes. The regulation of expression, the processing of the polycistronic primary transcript and the regulation of the RNA metabolism in the mitochondria of *Plasmodium* remains largely unknown.

Through high-throughput sequencing of cellular RNA, we discovered a population of small RNAs originating in the mitochondria of *P. falciparum*. Some of these short RNAs correspond to known mRNAs, however, a subset of short RNAs accumulates at the 5'-end of rRNA genes. Some of the cosRNAs were verified with RNA blots. Similar RNA accumulations can be detected in the organelles of higher plants and are caused by helical-hairpin repeat proteins like PPR proteins (pentatricopeptide repeat). These RNAs are termed clustered organellar short RNAs (cosRNAs).

To search for plant-like RNA binding proteins similar to PPR proteins we scanned the nuclear genome of *P. falciparum* for helical-hairpin repeat proteins. We found a novel protein family with repetitive helical elements of 37 amino acid length. Because of the helical-hairpin structure we termed these heptatricopeptide repeat (HPR) proteins. Targeting predictions of the 22 HPR proteins indicate almost exclusive mitochondrial localization. In the rodent Malaria parasite *P. berghei*, the mitochondrial localization for 7 HPR-Proteins was verified. In knockout studies, we also showed that almost all HPR proteins are essential for blood stages of *P. berghei*. In RNA-binding assays, one recombinant HPR protein showed unspecific interaction with mitochondrial transcripts but not with DNA. By broadening the search, we discovered that HPR proteins are found in multiple eukaryotic taxa. This

suggests an occurrence of the HPR proteins early in evolution. In humans, 6 proteins were identified containing HPR repeats. All those proteins belong to the FASTK protein family. The FASTK proteins localize to the mitochondria and play a role in the RNA metabolism of the human mitochondria.

# Inhaltsverzeichnis

<b>Zusammenfassung.....</b>	<b>ii</b>
<b>Summary .....</b>	<b>iv</b>
<b>Inhaltsverzeichnis .....</b>	<b>vi</b>
<b>1. Einleitung .....</b>	<b>1</b>
1.1 Verbreitung von Malaria und die Herausforderungen im Kampf gegen diese Krankheit .....	1
1.2 Zellstruktur und Lebenszyklus von <i>Plasmodium</i> .....	2
1.2.1 Zellstruktur und definitorische Bestandteile der Apicomplexa .....	2
1.2.2 Der komplexe Lebenszyklus von <i>P. falciparum</i> findet in zwei Organismen statt .....	3
1.3 Apicoplast und Mitochondrium der Apicomplexa zeigen prokaryotische Eigenschaften .....	5
1.3.1 Funktion und therapeutische Bedeutung des Apicoplasten .....	6
1.3.2 Expression des Apicoplastengenoms ist abhängig von kernkodierten Faktoren .....	7
1.3.3 Organisation und Funktion des Mitochondriums .....	9
1.3.4 Struktur und Expression des mitochondrialen Genoms .....	10
1.4 PPR- und OPR-Proteine sind dominante Faktoren in der organellären RNA Prozessierung und Stabilisierung in photoautotrophen Organismen .....	15
1.4.1 Struktur und Funktion von <i>Pentatricopeptide repeat</i> Proteinen (PPR) .....	15
1.4.2 Struktur und Funktion von <i>Octatricopeptide repeat</i> Proteinen (OPR) .....	18
1.4.3 Helikale Proteinfamilien spielen in den Organellen vieler Organismen eine wichtige Rolle .....	20
1.5 Kurze RNAs in den Organellen höherer Pflanzen können von PPR-Proteinen generiert werden .....	21
1.6 Zielstellung .....	23
<b>2. Material und Methoden .....</b>	<b>25</b>
2.1 Materialien .....	25
2.1.1 Parasiten .....	25
2.1.2 Mäuse .....	25
2.1.3 Chemikalien .....	25
2.1.4 Enzyme .....	26
2.1.5 Bakterienstämme .....	26
2.1.6 Plasmide .....	26
2.1.7 Antikörper .....	26
2.1.8 Oligonukleotide .....	27
2.1.9 Software .....	27
2.1.10 Datenbanken .....	27
2.2 Methoden .....	27

## Inhaltsverzeichnis

---

2.2.1	Inaktivierung von biologischem Material .....	27
2.2.2	Kultivierung von <i>Escherichia coli</i> .....	28
2.2.2.1	Kultivierung auf Agar-Agar-LB Platten .....	28
2.2.2.2	Kultivierung in flüssigem LB-medium .....	28
2.2.3	Kultivierung von <i>Plasmodium falciparum</i> .....	28
2.2.4	Amplifikation von DNA-Molekülen mittels PCR .....	29
2.2.5	Transformation von chemisch kompetenten <i>Escherichia coli</i> .....	30
2.2.6	Kolonie-PCR .....	30
2.2.7	Gelelektrophorese von Nukleinsäuren .....	31
2.2.8	Spektroskopische Analyse von Nukleinsäuren .....	31
2.2.9	Isolation und Aufreinigung von PCR Produkten/Plasmiden .....	32
2.2.10	Verwendung von T4-DNA Ligase .....	32
2.2.11	Verwendung von Restriktionsendonukleasen .....	32
2.2.12	Extraktion von RNA aus organischem Material .....	32
2.2.13	Extraktion von Nukleinsäuren mit Phenol/Chloroform .....	33
2.2.14	Präzipitation von Nukleinsäuren mit Ethanol .....	33
2.2.15	<i>In vitro</i> Transkription .....	34
2.2.16	Denaturierende RNA-Gelelektrophorese .....	34
2.2.17	Transfer von RNA auf eine Nylonmembran ( <i>Northern-blot</i> ) .....	35
2.2.18	Markierung, Hybridisierung und Detektion von radioaktiven Sonden .....	36
2.2.18.1	Radioaktive Markierung von DNA-Oligonukleotiden .....	36
2.2.18.2	Hybridisierung und Detektion der Sonden .....	36
2.2.19	Denaturierende RNA Urea-Polyacrylamid-Gelelektrophorese (Urea-PAGE) .....	36
2.2.20	Transfer von kurzen RNAs aus der Urea-PAGE auf Nylonmembranen .....	37
2.2.21	Aufreinigung von Erythrozyten aus Mäuseblut .....	38
2.2.22	Isolation genomischer DNA aus Blutstadien von <i>P. berghei</i> .....	38
2.2.23	Anreicherung von Organellen aus <i>P. falciparum</i> .....	39
2.2.24	Erstellung von cDNA-Bibliotheken für Hochdurchsatzsequenzierung .....	40
2.2.25	Bioinformatische Analyse von Hochdurchsatzsequenzierungsdaten .....	40
2.2.26	Quantifizierung von Proteinlösungen .....	41
2.2.27	Polyacrylamid-Gelelektrophorese (SDS-PAGE) .....	41
2.2.28	Proteintransfer auf Nitrocellulose-Membranen ( <i>Western-blot</i> ) .....	42
2.2.29	Heterologe Expression von Proteinen .....	43
2.2.29.1	Erstellung der Expressionskonstrukte .....	43
2.2.29.2	Expression der rekombinanten Proteine in <i>E. coli</i> .....	44
2.2.29.3	Reinigung von rekombinanten Fusionsproteinen mittels Affinitätschromatographie .....	45
2.2.30	RNA-Bindungsstudie von rekombinanten Proteinen .....	45
2.2.31	Giemsa Färbung für die Bestimmung der Parasitämie .....	46

2.2.32	Isolation von Parasiten aus infiziertem Blut.....	46
2.2.33	Transfektion von <i>Plasmodium berghei</i> ANKA .....	47
2.2.33.1	Anreicherung von Schizonten für die Transfektion.....	47
2.2.33.2	Isolation der Schizonten .....	48
2.2.33.3	Transfektion von <i>P. berghei</i> mittels Elektroporation.....	48
2.2.34	Fluoreszenzmarkierung von Proteinen in <i>P.berghei</i> .....	49
2.2.35	Deletion von Proteinen in <i>P. berghei</i> .....	50
2.2.36	Fluoreszenzmikroskopie .....	51
2.2.37	Reverse Transkription.....	51
2.2.38	Quantitative PCR .....	51
<b>3.</b>	<b>Ergebnisse .....</b>	<b>53</b>
3.1	Vorbemerkungen .....	53
3.2	Sequenzierung von organellären kurzen RNAs aus <i>P. falciparum</i> .....	53
3.2.1	Die Organellen von <i>P. falciparum</i> können durch differentielle Zentrifugation angereichert werden .....	54
3.2.2	Erstellung von Sequenzierbibliotheken und Hochdurchsatzsequenzierung kurzer RNAs aus <i>P. falciparum</i> .....	56
3.2.3	CosRNAs im Apicoplasten liegen innerhalb von tRNA und rRNA Sequenzen .....	57
3.2.4	Sequenzierung verbessert die Annotation der rRNA-Fragmente des Mitochondriums .....	58
3.2.5	Mitochondriale cosRNAs passen zu den 5'-Enden von beschriebenen rRNA- Fragmente .....	61
3.2.6	Einige mitochondriale cosRNAs können in RNA-Hybridisierungen detektiert werden .....	62
3.2.7	Einige mitochondriale cosRNAs definieren die Transkriptenden von mRNAs, andere liegen in intergenischen Bereichen .....	64
3.3	Identifikation einer neuen Familie von helikalen Proteinen in <i>Plasmodium</i> .....	67
3.3.1	Genomweite Suche mit einem PPR-Konsensus Motiv identifiziert keine neuen Proteine.....	67
3.3.2	Suche nach OPR-ähnlichen Proteinen identifiziert ein neues repetitives Element in <i>P. falciparum</i> Proteinen .....	68
3.3.3	Die Analyse der HPR-Proteine in <i>P. falciparum</i> deutet auf eine fast exklusiv mitochondriale Lokalisation hin .....	70
3.3.4	HPR-Proteine konnten auch in <i>P. berghei</i> identifiziert werden .....	72
3.4	HPR-Proteine außerhalb von <i>Plasmodium</i> .....	74
3.4.1	HPR-Proteine werden in fast allen untersuchten Alveolata identifiziert .....	75
3.4.2	HPR-Proteine können in vielen Taxa identifiziert werden, im Menschen werden FASTK-Proteine als HPR-Protein identifiziert .....	77
3.5	Funktionelle Untersuchung der neu entdeckten HPR-Proteine .....	79

3.5.1	Die HPR-Proteine sind dominiert durch helikale Elemente, die auch außerhalb der HPR-Motive liegen .....	79
3.5.2	Ausgewählte HPR-Proteine lokalisieren <i>in vivo</i> in Mitochondrien von <i>P. berghei</i> ..	81
3.6	Deletion von HPR-Proteinen deuten auf essentielle Funktion der HPR-Proteine in den Blutstadien hin .....	84
3.7	Analyse der RNA-Bindung von PbHPR1 und PbHPR2 <i>in vitro</i> .....	86
3.7.1	Expression von rekombinantem PbHPR1 und PbHPR2 in <i>E. coli</i> .....	87
3.7.2	Herstellung der mitochondrialen RNAs durch <i>in vitro</i> Transkription .....	90
3.7.3	Rekombinantes PbHPR1-Protein bindet <i>in vitro</i> unspezifisch mitochondriale RNAs .....	90
3.7.4	Rekombinante HPR-Proteine binden <i>in vitro</i> keine DNA .....	94
<b>4.</b>	<b>Diskussion .....</b>	<b>96</b>
4.1	Identifikation von cosRNAs in den Organellen von <i>P. falciparum</i> als potentielle Interaktionsstellen mit Proteinen .....	97
4.1.1	Die fehlenden cosRNAs im Apicoplasten deuten auf einen PPR-unabhängigen Mechanismus hin .....	97
4.1.2	In den Mitochondrien von <i>P. falciparum</i> akkumulieren cosRNAs an den Transkriptenden .....	98
4.1.3	Die 5'-Enden von cosRNAs überlappen mit den 5'-Enden von rRNA-Fragmenten .....	100
4.2	HPR-Proteine – eine neue Familie von helikalen Proteinen im RNA-Metabolismus der Mitochondrien .....	103
4.2.1	HPR-Proteine sind strukturell verwandt mit OPR-Proteinen, stellen jedoch eine eigene Familie dar .....	103
4.2.2	HPR-Proteine sind vor allem in den meisten Alveolata weit verbreitet .....	104
4.2.3	Besteht ein Zusammenhang zwischen HPR-Proteinen und fragmentierten rRNA-Genen im mitochondrialen Genom? .....	105
4.2.4	Die Deletion von HPR-Proteinen führt zu einem dramatischen Fitnessverlust in <i>P. berghei</i> .....	106
4.2.5	Welche Faktoren spielen für den Fitnessverlust der Nullmutanten von HPR-Genen eine Rolle? .....	107
4.2.6	HPR-Proteine im Menschen haben eine Funktion im RNA Metabolismus der Mitochondrien .....	109
4.2.7	HPR-Proteine binden <i>in vitro</i> unspezifisch RNA .....	111
4.2.8	Hypothetisches Modell für die Funktion von HPR-Proteinen im Mitochondrium ..	112
4.2.9	Eignen sich die HPR-Proteine für die Entwicklung neuer Medikamente gegen Malaria? .....	117
	<b>Literaturverzeichnis .....</b>	<b>119</b>
	<b>Anhang .....</b>	<b>138</b>



## Inhaltsverzeichnis

---

<b>Abkürzungen.....</b>	<b>152</b>
<b>Danksagung .....</b>	<b>154</b>
<b>Lebenslauf.....</b>	<b>Fehler! Textmarke nicht definiert.</b>
<b>Selbstständigkeitserklärung .....</b>	<b>156</b>

# 1. Einleitung

## 1.1 Verbreitung von Malaria und die Herausforderungen im Kampf gegen diese Krankheit

Nach den Zahlen der WHO (World Health Organization) wurden im Jahr 2016 über 216 Millionen Menschen weltweit mit Malaria infiziert, wobei ca. 445.000 Infektionen tödlich verliefen. Vor allem in Afrika stellt Malaria ein großes Gesundheitsproblem dar. Hier werden 90% der Todesfälle gezählt und vor allem Kinder unter 5 Jahren sind besonders stark betroffen (WHO 2017). Malaria bei Menschen wird von fünf einzelligen Parasiten der Gattung *Plasmodium* verursacht (*P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*, *P. knowlesi*), wobei *P. falciparum* für die meisten der Todesfälle verantwortlich ist.

Obwohl viele Ressourcen (etwa 2,7 Mrd. US-Dollar weltweit, WHO 2017) für die Bekämpfung von Malaria und die Erforschung neuer Medikamente bereitgestellt werden, stellt die Entstehung von neuen Resistenzen eine große Herausforderung dar. Ein Beispiel ist das Medikament Chloroquin, welches Mitte des 20. Jahrhunderts gegen Malaria eingesetzt wurde, heute jedoch kaum verwendet werden kann, da eine breitflächige Resistenz von *P. falciparum* eingetreten ist (Payne 1987, Wellems und Plowe 2001). Auch andere Medikamente wie beispielsweise Sulfadoxin/Pyrimethamin verloren ihre Wirkung durch Resistenzentwicklung (Antony und Parija 2016, Petersen et al. 2011). Die heute von der WHO empfohlene Therapie ist Artemisinin-basiert und wird als ACT (*artemisinin-based combination therapies*) bezeichnet. Zu den Medikamenten, die bei der ACT in Kombination mit Artemisinin eingesetzt werden gehören unter anderem Mefloquin oder Sulfadoxin/Pyrimethamin, die jeweilige Kombination wird basierend auf den lokalen Gegebenheiten variiert (Nosten und White 2007). Obwohl heute noch effizient, zeigen neuere Studien die Gefahr einer entstehenden Resistenz gegen Artemisinin, unter anderem in Südostasien auf (Fairhurst et al. 2012, Imwong et al. 2017). Neben der chemischen Therapie sind auch andere Methoden, wie beispielsweise der Einsatz von Moskitonetzen, ein wichtiger Bestandteil der Anstrengungen Malaria einzudämmen. Es wird vermutet, dass bereits der Einsatz

dieser Netze alleine eine Vielzahl von Infektionen verhindert hat (Bhatt et al. 2015). Jedoch bleibt die Entwicklung neuer Medikamente essentiell für den Kampf gegen Malaria. Weiterhin ist auch die grundlegende Erforschung und das bessere Verständnis dieser Parasiten ein wichtiger Baustein zur erfolgreichen Bekämpfung von *Plasmodium*, da durch die Grundlagenforschung neue Zellbestandteile und Abläufe entdeckt werden können, die potentielle Ansatzpunkte für die Medikamentenentwicklung darstellen.

## 1.2 Zellstruktur und Lebenszyklus von *Plasmodium*

### 1.2.1 Zellstruktur und definitorische Bestandteile der Apicomplexa

Die Parasiten der Gattung *Plasmodium* gehören zur Gruppe der Apicomplexa, einem Phylum von einzelligen eukaryotischen Parasiten, welche

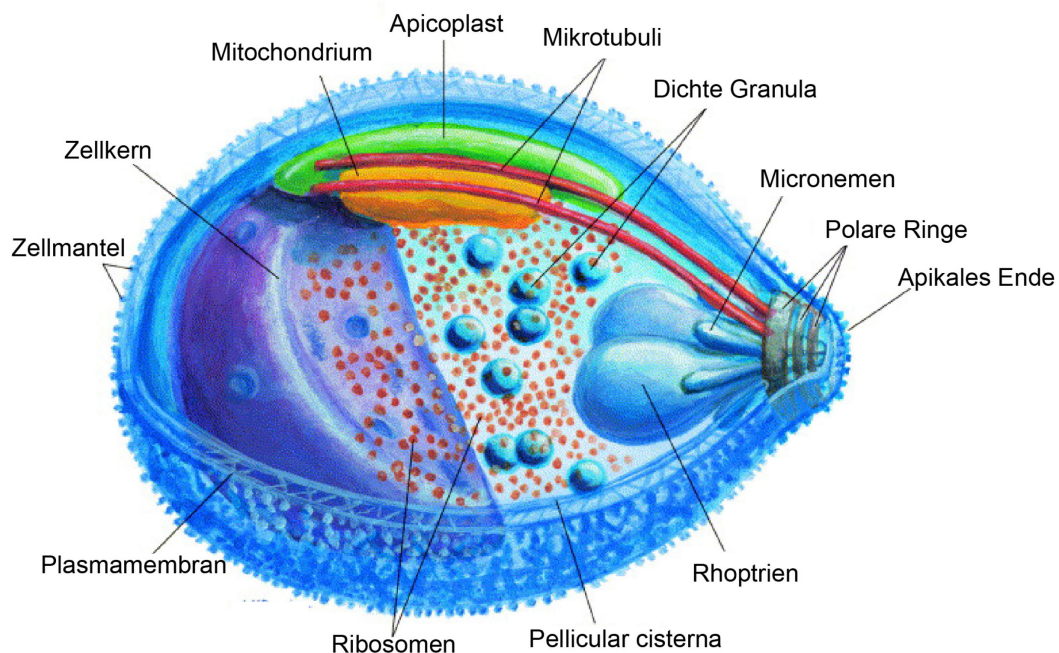


Abbildung 1: Zellstruktur von *P. falciparum*. Die Zellen von Plasmodium haben eine polare Zellstruktur mit einem apikalen Ende. Wie die meisten Eukaryoten besitzen die Parasiten klassische Zellkomponenten (bspw. Zellkern, Mikrotubuli, u.Ä.). Als Besonderheit sind die sekretorischen Organellen (Micronemen, Dichte Granula und Rhoptrien) zu nennen. Außerdem beinhaltet jede Zelle ein einzelnes Mitochondrium und ebenfalls nur einen nicht-photosynthetischen Apicoplasten. (Verändert nach Bannister et al. 2000).

zusammen mit beispielsweise Dinoflagellaten und Wimpertierchen das Taxon Alveolata bilden. Die Zelle von *Plasmodium spp.*, wie auch die meisten Apicomplexa, besitzen einige einzigartige Organellen und Bestandteile. Die schematische Zellstruktur von *P. falciparum* ist in Abbildung 1 gezeigt. Die Zellen von *Plasmodium* haben durch das definierte apikale Ende eine polare Struktur. Neben den klassischen Zellbestandteilen (Zellkern, Golgi, Endoplasmatisches Retikulum, Mikrotubuli u.a.) sind auch spezielle, sekretorische Organellen zu finden (Micronemen und Rhoptrien), welche Proteine sekretieren und während des Invasionsprozesses wichtig sind (zusammengefasst in Dubremetz et al. 1998). Für die Ziele dieser Arbeit besonders bedeutsam und interessant sind zwei Organellen, das Mitochondrium und der photosynthetisch nicht aktive Plastid – der Apicoplast. Auf diese beiden Organellen wird in 1.3 detaillierter eingegangen.

### **1.2.2 Der komplexe Lebenszyklus von *P. falciparum* findet in zwei Organismen statt**

Der Lebenszyklus von *Plasmodium* (hier am Beispiel von *P. falciparum*) ist komplex und verläuft zum einen Teil in einem Vertebraten (*Homo sapiens*) und zum anderen Teil in einem Insekt (weibliche Stechmücken der Gattung *Anopheles*). In Abbildung 2 ist eine schematische Darstellung des Lebenszyklus zu finden. Die Infektion der Menschen findet durch den Stich eines infektiösen Moskitos statt, hierbei werden *P. falciparum* Sporozoiten über den Speichel in die Haut transferiert (Sinnis und Zavala 2012). Von der Stichstelle aus gleiten die Parasiten aktiv bis ein Blutgefäß erreicht wird. Durch den Blutstrom gelangen die Sporozoiten zur Leber (Douglas et al. 2015). Von den Blutgefäßen der Leber (Sinusoid) wandern die Parasiten durch verschiedene Zellschichten in die Hepatozyten, wo eine parasitophore Vakuole gebildet wird (Mota et al. 2001). In diesem Stadium verläuft die Infektion klinisch unsichtbar, es werden jedoch durch Teilung zwischen 10.000 und 30.000 neue Tochterzellen oder Merozoiten gebildet (Prudêncio et al. 2006). Diese werden schließlich in den Blutstrom entlassen und es beginnt der im Menschen klinisch relevante Teil des Lebenszyklus in der Blutbahn.

Die freigesetzten Merozoiten dringen in die Erythrozyten ein und der Parasit durchläuft mehrere Stadien des Wachstums. Dabei entwickeln sich aus dem anfänglichen Ringstadium die Trophozoiten. Zur Energiegewinnung wird Glukose

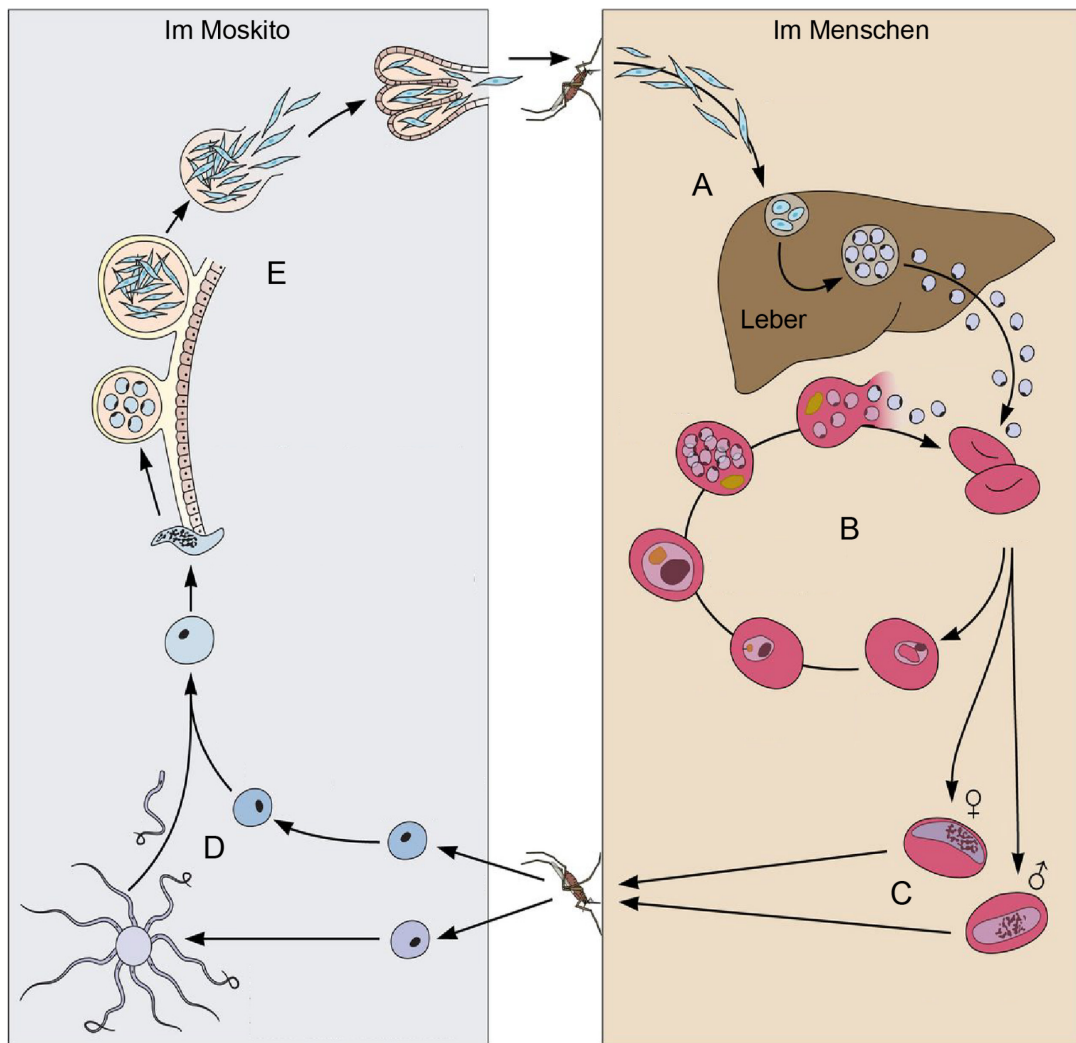


Abbildung 2: Lebenszyklus von *P. falciparum*. Im Lebenszyklus wechselt der Parasit zwischen zwei Wirten. Die wichtigsten Abschnitte sind hier schematisch aufgeführt. **A)** Durch den Biss eines Moskitos werden Sporozoiten auf den Menschen übertragen, diese wandern dann zur Leber, infizieren Hepatozyten und vermehren sich zu einer großen Menge Merozoiten. **B)** Diese werden in die Blutbahn freigesetzt, invadieren Erythrozyten und replizieren Tochterzellen. **C)** Einige Zellen gehen eine irreversibel sexuelle Differenzierung ein und es entstehen Gametozyten, welche durch Blutmahlzeiten von Moskitos aufgenommen werden. **D)** Ein Mikro- und ein Makrogamet fusionieren zu einer diploiden Zygote. **E)** Nach der Umwandlung in die Ookinete wandert diese durch das Epihel und haftet sich als Oozyste an die Basallamina des Darms der Insekten. Nach einer 24-stündigen Reifungsperiode werden neue, haploide Sporozoiten asexuell gebildet. Diese wandern in die Speicheldrüsen des Moskitos und können bei einer Blutmahlzeit wieder auf einen Menschen übertragen werden (verändert nach Lee et al. 2014).

aus der Wirtszelle genutzt. Zusätzlich liefert auch der Abbau des Hämoglobins die nötigen Ressourcen für den Parasiten (Aikawa et al. 1966, Elliott et al. 2008). Die

Trophozoiten entwickeln sich dann in Schizonten, welche sich in 8 bis 36 Merozoiten teilen und anschließend die Blutzelle lysieren (Glushakova et al. 2010). Ein kleiner Teil der intrazellulären Parasiten entwickelt sich nicht zu Merozoiten, sondern folgt einer sexuellen Differenzierung hin zu den Gametozyten. Dieses Stadium kann von Moskitos aufgenommen werden und es folgt die sexuelle Entwicklung (Kooij und Matuschewski 2007).

Bei der Aufnahme der Gametozyten durch den Mosquito werden die Gameten gebildet. Die Mikrogametozyten bilden dabei 8 Plasmafortsätze aus die jeweils einen Kern tragen und sich anschließend abschnüren. Dieser Prozess wird als Exflagellation bezeichnet. Ein Mikro- und ein Makrogamet verschmelzen zu einer Zygote, einem diploiden Stadium, welche sich innerhalb von 24 Stunden in eine Ookinete verwandelt. Diese wandert durch das Epithelium des Verdauungstrakts und bildet an der Basallamina eine Oozyste und beginnt, asexuell Sporozoiten zu replizieren. Diese werden bei der Lyse der Zyste freigesetzt und wandern in die Speicheldrüsen des Moskitos, wo durch einen Stich die erneute Infektion eines Menschen stattfinden kann (Zusammengefasst in Cowman et al. 2016).

### **1.3 Apicoplast und Mitochondrium der Apicomplexa zeigen prokaryotische Eigenschaften**

Während des Blutzyklus von *Plasmodium* durchlaufen sowohl der Apicoplast als auch das Mitochondrium morphologische Änderungen. In den frühen Stadien des Blutzyklus (Merozoiten, den Ringstadien und den Trophozoiten) ist sowohl der Apicoplast als auch das Mitochondrium als eine längliche, wurmartige Struktur erkennbar. In den Stufen der Schizogonie verästelt und verlängern sich das Mitochondrium und der Apicoplast (van Dooren et al. 2005). Durch mikroskopische Untersuchungen wurde gezeigt, dass Apicoplast und Mitochondrium zumindest in den Blutstadien in der Zelle direkt nebeneinander liegen und einzelne Kontaktpunkte besitzen (Hopkins et al. 1999). Es ist möglich, dass diese räumliche Nähe dem Austausch von Metaboliten dient und so die beiden Organellen nicht nur morphologisch, sondern auch funktionell verbunden sind (Ralph et al. 2004). Der Apicoplast und das Mitochondrium sind für den Parasiten essentiell und aufgrund

des prokaryotischen Ursprungs eröffnen sie spannende Ansätze für die klinische und die grundlegende Forschung.

### 1.3.1 Funktion und therapeutische Bedeutung des Apicoplasten

Durch die Entdeckung des Apicoplasten zeigte sich, dass apicomplexe Parasiten ein Organell besitzen, welches eher für Algen und Pflanzen definierend ist. Während in den Chloroplasten von phototropen Organismen die Photosynthese stattfindet, ist diese Funktion bei den Apicomplexa verloren gegangen (McFadden et al. 1996). Die Entstehung des Apicoplasten geht auf die sekundäre Endosymbiose der apicomplexen Vorläufer mit einer Rotalge zurück (Lim und McFadden 2010). Diese Abstammung macht den Apicoplasten zu einem attraktiven Ziel der Medikamentenentwicklung. Eine Inhibition der Translation bzw. Transkription des Apicoplasten führt in *P. falciparum* zu einem Phänotyp, der als verzögerter Zelltod beschrieben wird. Die Behandlung von Blutstadien von *P. falciparum* mit Doxycyclin, einem Inhibitor für prokaryotische Translation, führt in den ersten 48 Stunden der Behandlung nur zu einem geringen Effekt (Dahl et al. 2006). In dieser Zeitspanne durchläuft der Parasit einen Zyklus im Blut und die gebildeten Merozoiten können in neue Erythrozyten eindringen. In dem darauffolgenden, zweiten Blutzyklus ist der essentielle Proteinimport von kernkodierten Proteinen in den Apicoplasten jedoch gestört und auch die Replikation des 35 kb langen Genoms kommt zum Erliegen (Dahl et al. 2006). Dies führt schließlich dazu, dass der Apicoplast während der Schizogonie nicht repliziert werden kann. Die Parasiten können dadurch die Schizogonie nicht vollenden und es findet keine Lyse des Erythrozyten statt (Dahl und Rosenthal 2007).

Das Kerngenom von *P. falciparum* wurde vollständig sequenziert und beinhaltet auf 14 Chromosomen etwa 23 Megabasen mit ca. 5.500 kodierten Proteinen (Gardner et al. 2002). Bei der Genomsequenzierung wurde ein 35 kb langes von den Chromosomen des Kerns unabhängiges DNA Fragment identifiziert, welches dem Apicoplasten zugeordnet werden konnte (McFadden et al. 1996). Neben ribosomalen Proteinen und Struktur-RNAs (transfer-RNAs und ribosomale-RNAs) wird auch eine RNA-Polymerase, das Chaperon ClpC für die Proteinfaltung und mit SufB eine Untereinheit für den Syntheseweg von Eisen-Schwefel-Verbindungen auf dem Genom kodiert (Wilson et al. 1996). Für die Expression des

apicoplastidären Genoms sind Proteine des Kerngenoms nötig. Es werden ca. 400 kernkodierte Proteine im Zytosol translatiert und post-translational in den Apicoplasten transportiert (Foth et al. 2003). Der Proteintransport wird durch ein Amino-terminales Signalpeptid initiiert. Während des Imports müssen die Proteine die vier Hüllmembranen des Apicoplasten überwinden. Die inneren beiden Membranen werden hierbei scheinbar wie bei Pflanzen durch die beiden Proteinkomplexe TiC und ToC (*Translocons of the Inner bzw. Outer Chloroplast membranes*) überwunden (Glaser et al. 2012, Soll und Schleiff 2004). Wie die beiden äußeren Membranen überbrückt werden, ist noch nicht abschließend geklärt. Neue Studien lassen aber vermuten, dass ein ERAD basiertes System (*ER-associated protein degradation pathway*) für den Transport über die Membranen verantwortlich sein könnte (Agrawal et al. 2013, Spork et al. 2009).

Anhand der importierten Proteine und des genetischen Potentials des Apicoplastengenoms konnte festgestellt werden, dass der Apicoplast der Reaktionsraum für einige Stoffwechselwege ist. Unter diese fällt beispielsweise die Fettsäuresynthese (FAS-II, *Fatty acid synthesis pathway type 2*), die Synthese von Eisen-Schwefelverbindungen zusammen mit dem Mitochondrium sowie die Herstellung von Isoprenoid-Vorläufern (Ralph et al. 2004). Die Inhibition des Apicoplasten durch Antibiotika ist für den Parasiten tödlich, jedoch konnten die Parasiten durch die exogene Zugabe von Isopentenylpyrophosphat (IPP) gerettet werden. Dies deutet darauf hin, dass dies der einzige apicoplastidäre Stoffwechselweg ist, der in diesem Stadium eine essentielle Rolle spielt (Yeh und DeRisi 2011). In den Moskitostadien von *P. berghei* konnte gezeigt werden, dass die Herstellung von Eisen-Schwefelverbindungen essentiell für die Replikation von Sporozoiten ist (Charan et al. 2017). Es zeigt sich also, dass der Apicoplast nicht nur morphologisch, sondern auch funktional eine Änderung im Lebenszyklus durchläuft.

### **1.3.2 Expression des Apicoplastengenoms ist abhängig von kernkodierten Faktoren**

Anhand der Mechanismen und Faktoren der Genexpression des Apicoplasten lässt sich die prokaryotische Herkunft dieses Organells erkennen. Es konnte gezeigt werden, dass Antibiotika, welche bakterielle 70S Ribosomen



inhibieren, wirksame Medikamente gegen Malaria sind. Dies lässt vermuten, dass die Translation des Apicoplasten essentiell für das Überleben des Parasiten ist. Auch Antibiotika wie Rifampicin, die inhibierend auf bakterielle, DNA-abhängige RNA-Polymerasen wirken, können als Medikament gegen Malaria eingesetzt werden (Goodman et al. 2007). Die Transkription der Apicoplastengene erfolgt in polycistronischen Einheiten, welche durch Prozessierung in monocistronische Transkripte geschnitten werden (Gardner et al. 1991, Nisbet et al. 2016). In den Plastiden von Pflanzen wird die Transkription von zwei Polymerasen durchgeführt. Diese beiden Enzyme werden als NEP (*Nuclear encoded Polymerase*) und PEP (*Plastid encoded Polymerase*) bezeichnet (zusammengefasst in Liere et al. 2011). Zwar wurden in Apicomplexa NEP-ähnliche Phagentyppolymerasen identifiziert, jedoch scheinen diese nur für die Transkription in Mitochondrien verantwortlich zu sein (Ke et al. 2012). Die im Apicoplasten kodierte Polymerase zeigt deutliche Ähnlichkeiten zu bakteriellen Polymerasen und besitzt ähnliche Untereinheiten wie die PEP, jedoch wurden bisher keine Sigma-Faktoren identifiziert, die in Chloroplasten für die Promotorerkennung verantwortlich sind. Gleichzeitig konnten im Apicoplastengenom keine konservierten Promotorsequenzen identifiziert werden, sodass unbekannt ist, wie die Initiation der Transkription erfolgt (Liere et al. 2011, Nisbet und McKenzie 2016).

Neben dem lückenhaften Verständnis der Transkription ist auch der genaue Mechanismus der Prozessierung der polycistronischen RNAs weitestgehend unbekannt. Es ist jedoch naheliegend, dass an diesem Prozess RNA-bindende Proteine beteiligt sind. Es wurden zwei Proteine der „*Pentatricopeptide repeat proteine*“ (PPR) Familie identifiziert. Mitglieder dieser Familie spielen in höheren Pflanzen eine große Rolle im RNA-Metabolismus der Organellen (auf die PPR-Proteine wird genauer in 1.4.1 eingegangen). Eines der PPR-Proteine ist im Apicoplasten lokalisiert und essentiell für den Parasiten (Hicks et al. 2018). Das Protein bindet *in vitro* spezifisch Transkripte des Apicoplasten und könnte ein Faktor für die Transkriptprozessierung sein (Hicks et al. 2018). Die identifizierten Prozessierungsstellen von Apicoplastentranskripten decken sich häufig mit den Transkriptenden von tRNA-Genen, sodass eine Prozessierung nach dem tRNA-*punctuation processing* Modell (Ojala et al. 1981) möglich scheint (Nisbet et al. 2016). In diesem Modell werden bei der tRNA-Reifung durch RNase P und RNase

Z gleichzeitig reife mRNAs und rRNAs erzeugt (zusammengefasst in Taanman 1999). Im Genom von *P. falciparum* wurde eine RNase P identifiziert, deren bioinformatische Analyse auf eine Lokalisation im Apicoplasten hindeutet (Lechner et al. 2015). Eine RNase Z wurde im Genom von *Plasmodium* bisher nicht beschrieben. Insgesamt sind bisher nur wenige potentielle Faktoren des apicoplastidären RNA Metabolismus beschrieben worden. Immerhin wurde in *P. vivax* ein potentiell RNA bindendes Protein mit Plastidenlokalisation identifiziert. Das Protein enthält ein RNA-Bindedomäne (RRM, *RNA-recognition motif*) und durch *in vitro* Experimente konnte die Bindung an Uracil-reiche RNAs gezeigt werden (García-Mauriño et al. 2017).

### 1.3.3 Organisation und Funktion des Mitochondriums

Das Mitochondrium von *Plasmodium spp.* ist von zwei Hüllmembranen umgeben. In Blutstadien besitzt jeder Parasit jeweils nur ein Mitochondrium. Eine der wichtigsten Funktionen von eukaryotischen Mitochondrien ist die Energiegewinnung aus dem Tricarbonsäurezyklus (TCA, *tricarboxylic acid cycle*), bei dem aus der Glykolyse stammendes Pyruvat genutzt wird, um sowohl Energie, als auch verschiedene zentrale Metaboliten bereitzustellen. Einer der Schlüsselkomplexe ist der Pyruvatdehydrogenasekomplex (PDH), ein Multienzymkomplex, der die oxidative Decarboxylierung von Pyruvat zu Acetyl-CoA katalysiert, welches als Substrat im TCA genutzt wird (zusammengefasst in Behal et al. 1993). In *Plasmodium* konnte nur ein vollständiges PDH-Enzym identifiziert werden, welches im Kern kodiert und in den Apicoplasten transportiert wird, wo es vermutlich an der Fettsäure-Synthese beteiligt ist und nicht primär am TCA (Foth et al. 2004, Ralph et al. 2004). Dies könnte einer der Gründe sein, warum Glukose zwar die Hauptenergiequelle für die Blutstadien von *Plasmodium spp.* ist, allerdings zum größten Teil in Laktat umgewandelt wird (Bryant et al. 1964). Obwohl zu den meisten weiteren Enzymen des TCA Homologe in *P. falciparum* identifiziert werden konnten (zusammengefasst in van Dooren et al. 2006), sind diese während der asexuellen Blutstadien nicht essentiell. In Deletionsmutanten wurden sechs der acht Enzyme deletiert, ohne das Wachstum der Blutstadien zu beeinflussen (Ke et al. 2015). Jedoch scheint in Moskitostadien die Funktion einiger Enzyme essentiell zu sein. Beispielsweise waren Mutanten ohne  $\alpha$ -Ketoglutarat-Dehydrogenase nicht in

der Lage in Moskitos eine Oozyste zu bilden (Ke et al. 2015). Anhand von Lebenszyklus-spezifischen Transkriptom- und Proteomdaten von *P. berghei* wurde vermutet, dass die Enzyme des TCA, der oxidativen Phosphorylierung und der Atmungskette in den Moskitostadien häufiger sind, was darauf hindeutet, dass hier die Mitochondrien als Energielieferant wichtiger sind als in den asexuellen Blutstadien (Hall et al. 2005).

Eine essentielle Funktion des Mitochondriums in Blutstadien scheint die Elektronentransportkette (ETC) zu sein. Hierbei spielt besonders die Regeneration von Ubichinon als Elektronenakzeptor in der Pyrimidin-Biosynthese eine entscheidende Rolle (Painter et al. 2007). An einigen Schritten der Elektronentransportkette sind Komplexe mit Häm-Gruppen beteiligt. Die letzten Schritte der *de novo* Synthese von Häm-Gruppen findet im Mitochondrium statt, jedoch finden die vorherigen Schritte des Syntheseweges sowohl in Mitochondrium, Zytosol und Apicoplast statt (Ke et al. 2014). In Blutstadien ist *P. falciparum* nicht auf die *de novo* Synthese von Häm-Gruppen angewiesen, vermutlich werden die Häm-Gruppen aus dem Erythrozyten genutzt (Ke et al. 2014). Diese Möglichkeit steht dem Parasiten in den Moskitostadien jedoch nicht zur Verfügung, sodass hier der Syntheseweg der Häm-Gruppen essentiell wird (Ke et al. 2014). Es ist zu erkennen, dass auch das Mitochondrium in den verschiedenen Stufen des Lebenszyklus eine unterschiedliche metabolische Bedeutung hat.

### 1.3.4 Struktur und Expression des mitochondrialen Genoms

Das mitochondriale Genom von *Plasmodium* gehört zu den kleinsten beschriebenen, mitochondrialen Genomen. Verglichen mit den Mitochondrien des Menschen und seinen 37 genkodierenden Abschnitten auf ~16,5 kb (zusammengefasst in Taanman 1999) ist der Gengehalt des *Plasmodium spp.* Genoms mit 3 proteinkodierenden Sequenzen (Cytochrom b, Cytochrom Oxidase I und III) auf ~6 kb stark reduziert (Feagin 1992). Das Genom liegt linear als Konkatamer vor (zwischen 30 und 100 Wiederholungen) und wird aktiv transkribiert (Vaidya und Arasu 1987). Die Karte des mitochondrialen Genoms ist in Abbildung 3 dargestellt. Die Gene liegen sehr dicht aneinander. Meist befinden sich zwischen den einzelnen kodierenden Abschnitten nur wenige Nukleotide (Feagin et al. 2012). Es ist daher unwahrscheinlich, dass Promotoren für einzelne Gene existieren. Die

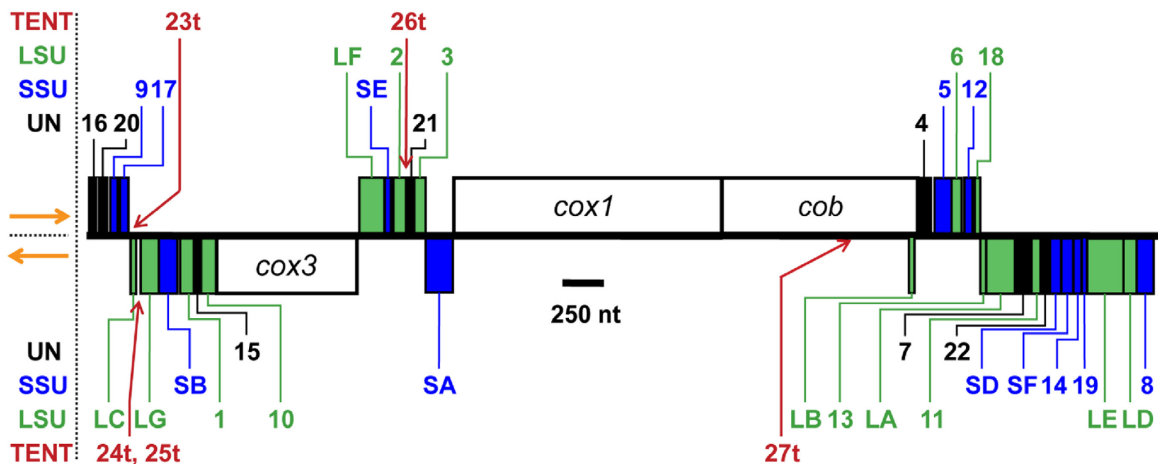


Abbildung 3: Karte des mitochondrialen Genoms. Das 6 kb lange Genom enthält neben den drei proteinkodierenden Genen (*coxI*, *coxIII* und *cob*) stark fragmentierte ribosomale RNAs (Zahlen bzw. S für die kleine Untereinheit, L für die große Untereinheit des Ribosoms). Fragmente, die bioinformatisch den Untereinheiten des Ribosoms zugeordnet wurden, sind in grün (große Untereinheit; LSU) und blau (kleine Untereinheit, SSU) dargestellt. Fragmente, die nicht zugeordnet werden konnten, sind in schwarz dargestellt; die Position von hypothetischen Fragmenten ist mit roten Pfeilen markiert (aus Feagin et al. 2012)

Vermutung, dass die Transkription in längeren, polycistronischen Abschnitten erfolgt, liegt nahe und konnte auch bestätigt werden.

Während das Genom von menschlichen Mitochondrien vollständige rRNA-Gene enthält (zusammengefasst in Taanman 1999), sind die rRNA Gene des apicomplexen Mitochondriums fragmentiert. Die Sequenz der Fragmente zeigt Ähnlichkeiten mit den RNAs der großen und der kleine Untereinheit von Ribosomen, jedoch sind die Fragmente über das Genom verstreut und teilweise auf unterschiedlichen Strängen kodiert (Feagin et al. 2012, Feagin et al. 1992, Feagin et al. 1997, siehe Abbildung 3). Die Länge der Fragmente variiert zwischen 20 und 200 Nukleotiden (nt). Insgesamt konnten 34 potentielle rRNA-Fragmente identifiziert werden, von denen in einer bioinformatischen Untersuchung, basierend auf Sequenzhomologie und wichtigen Strukturelementen, 27 Fragmente einer ribosomalen Untereinheit zugeordnet werden konnten (in Abbildung 3 grün und blau dargestellt, Feagin et al. 2012). Die verbleibenden sieben Fragmente konnten nicht zugeordnet werden (in Abbildung 3 in schwarz dargestellt). Zusätzlich zu den 34 beschriebenen Fragmenten wurden fünf weitere Fragmente vermutet, die in

seltenen cDNA Klonen und EST-Analysen auftraten (Raabe et al. 2010, Feagin et al. 2012, in Abbildung 3 in rot eingezeichnet). Ob diese Fragmente ebenfalls in den ribosomalen Untereinheiten vorliegen, ist nicht bekannt. Im Längenvergleich mit anderen mitochondrialen rRNAs scheinen die rRNAs von *P. falciparum* verkürzt. Rechnet man die verschiedenen rRNA Fragmente zusammen, ergeben sich für die große Untereinheit 1233 nt (Mensch: 1559 nt); für die kleine Untereinheit ergeben sich 804 nt (Mensch: 954 nt) (Feagin et al. 2012). Insgesamt lässt sich feststellen, dass die rRNAs und sehr wahrscheinlich auch die mitochondrialen Ribosomen von *Plasmodium* eine stark abgewandelte, einzigartige Struktur aufweisen. Wie die mitochondrialen Ribosomen von *Plasmodium* assembliert werden, ist vollkommen offen.

Die Expression des Organellengenoms in *Plasmodium spp.* ist stark abhängig von kernkodierten Proteinen, die post-translational in das Mitochondrium transferiert werden. Die Anzahl der importierten Proteine ist nicht genau bekannt. Durch Vorhersagen, mit speziell auf *P. falciparum* eingestellten Algorithmen, konnten zwischen 391 (7,1 % des Proteoms) und 1.177 (22 % des Proteoms) Proteine bestimmt werden, die potentiell ins Mitochondrium transportiert werden (Bender et al. 2003). Der Proteinimport findet vermutlich durch die ToM und TiM (*translocase of the outer membrane* bzw. *inner membrane*) Proteinkomplexe statt (van Dooren et al. 2006). Da das Genom des Mitochondriums keinerlei tRNAs, ribosomale Proteine oder andere Translationsfaktoren enthält, müssen alle diese Faktoren importiert werden. Sowohl in *P. falciparum*, als auch in *T. gondii* konnten zytosolische tRNAs im Mitochondrium nachgewiesen werden (Esseiva et al. 2004, Sharma und Sharma 2015). Zur Regeneration von tRNAs werden Aminoacyl-tRNA-Synthetasen (AaRS) benötigt. Diese Enzyme katalysieren die Verknüpfung von spezifischen Aminosäuren mit den entsprechenden tRNAs. Diese werden anschließend von den Ribosomen für die Proteinbiosynthese genutzt (zusammengefasst in Ibba und Söll 2000). In *P. falciparum* wurden bis jetzt 36 AaRS identifiziert, jedoch wurde nur für eine einzelne AaRS für Phenylalanin-tRNAs eine Lokalisation im Mitochondrium nachgewiesen (Sharma und Sharma 2015). Dies würde bedeuten, dass die tRNAs im Mitochondrium von *P. falciparum* einmalig verwendet und anschließend abgebaut werden. Ob tRNAs aus dem Mitochondrium

exportiert werden, um beispielsweise im Zytosol regeneriert zu werden, ist unbekannt.

Unter den importierten Proteinen findet sich auch die für die Transkription nötige Polymerase. Hierbei handelt es sich um eine Polymerase mit Ähnlichkeit zu den Polymerasen von Phagen (T3/T7 Phagen). Es konnten keine weiteren potentiellen Polymerasen identifiziert werden (Li et al. 2001). Eine Deletionsmutante dieses Proteins ist nicht lebensfähig (Ke et al. 2012), was zeigt, dass die Expression des Genoms bzw. mindestens eins Gens des Genoms in asexuellen Blutstadien essentiell ist. Es konnten Transkripte von bis zu 6 kb Länge nachgewiesen werden, was vermuten lässt, dass das Genom polycistronisch transkribiert und prozessiert wird (Ji et al. 1996). In *Northern*-Hybridisierungen konnten die drei proteinkodierenden Transkripte als monocistronische Formen nachgewiesen werden (Feagin und Drew 1995). Auch die meisten kurzen rRNA Fragmente wurden in *Northern*-Hybridisierungen als monocistronische Transkripte detektiert (Feagin et al. 2012). Über den Mechanismus dieser Prozessierung und über die beteiligten Faktoren ist wenig bekannt. Neben dem PPR Protein des Apicoplasten (siehe 1.3.2) konnte ein weiteres PPR Protein identifiziert werden, dessen Lokalisation unbekannt ist, das aber möglicherweise in die Mitochondrien importiert wird (Hicks et al. 2018). Eine Prozessierung nach dem aus humanen Mitochondrien bekannten tRNA-*punctuation processing* Modell, bei dem RNAsen durch die Prozessierung von tRNAs aus polycistronischen Vorläufern gleichzeitig reife mRNA und rRNA generieren, scheint unwahrscheinlich, da wie oben erwähnt keine tRNAs auf dem Genom vorliegen (Ojala et al. 1981).

Ein weiteres interessantes Detail ist, dass die drei mRNAs des Mitochondriums am 3'-Ende polyadenyliert werden, jedoch ist die Länge der Poly-A-Schwänze je nach mRNA unterschiedlich. So wurden bis zu 10 nt am 3'-Ende der *coxI* mRNA angehängt, während die *coxIII*-RNA keine bzw. nicht detektierbare Anhänge besitzt (Rehkopf et al. 2000). In den Mitochondrien anderer Organismen kann die Polyadenylierung von Transkripten unterschiedliche Auswirkungen haben. In pflanzlichen Mitochondrien ist die Polyadenylierung von Transkripten ein Signal zum Abbau dieser RNA (zusammengefasst in Gagliardi et al. 2004). Im Gegensatz dazu ist die Polyadenylierung von Transkripten in humanen Mitochondrien nötig für

die Generierung des Stopp-Codons UAA; gleichzeitig konnte gezeigt werden das polyadenylierte Transkripte eine erhöhte Stabilität aufweisen (zusammengefasst in Gagliardi et al. 2004). Ob die Polyadenylierungen der Transkripte des apicomplexen Mitochondrien eine funktionale Rolle in der Prozessierung oder Stabilisierung der RNAs spielt, ist unbekannt (Rehkopf et al. 2000).

In Ribosomen sind neben den rRNAs eine Vielzahl von ribosomalen Proteinen beteiligt. In *Plasmodium spp.* wurden ribosomale Proteine bioinformatisch (basierend auf Sequenzhomologie) identifiziert und die potentielle Lokalisation im Mitochondrium untersucht (Gupta et al. 2014). Einige Proteine, die in vergleichbaren Ribosomen essentiell sind, scheinen jedoch in *Plasmodium spp.* nicht vorhanden zu sein, sodass die Proteinzusammensetzung der mitochondrialen Ribosomen als abgeleitet betrachtet werden kann (Gupta et al. 2014). Es ist anzunehmen, dass aus den derzeit bekannten rRNA-Fragmenten und den importierten ribosomalen Proteinen ein funktionsfähiges Ribosom gebildet wird, welches die Translation der drei proteinkodierenden Gene katalysiert. Einen direkten Nachweis der Translation im Mitochondrium gibt es derzeit nicht. Ein Hinweis auf die Existenz der mitochondrialen Translation ist die Resistenz gegen das Medikament Atovaquon. Zwei Punktmutationen im mitochondrialen *cob* Gen wurden beschrieben, die zur Resistenz der Parasiten gegen Atovaquon führen (Korsinczky et al. 2000). Die Punktmutationen führen zu einem einzelnen oder doppelten Aminosäureaustausch (Korsinczky et al. 2000). Die einfachste Erklärung für diesen Befund ist, dass die durch die Punktmutation veränderte Proteinsequenz des Cytochrom B die Empfindlichkeit der mitochondrialen ETC für Atovaquon verändert. Folglich müssen die mutierten Proteinversionen translatiert werden. Dies lässt auf translationsaktive mitochondriale Ribosomen schließen. Das mitochondrial lokalisierte, ribosomale Protein L13 in *P. falciparum* wurde durch Mutanten untersucht, die induzierbar die Proteinmengen des L13 Proteins reduzieren (Ke et al. 2018). Die Induktion der Mutante führt zur Reduktion des mitochondrialen Membranpotentials, zur Inhibition der ETC und schließlich zum Tod des Parasiten (Ke et al. 2018). Dies ist ein weiterer Hinweis auf die essentielle Rolle der Translation der drei mitochondrialen Proteine durch mitochondriale Ribosomen in *P. falciparum*.

Es ist somit zu erkennen, dass das Mitochondrium eine sehr komplexe Expressionsmaschinerie besitzt, die kaum studiert wurde und scheinbar ausschließlich aus kernkodierten Faktoren besteht. Die Störung der RNA Prozessierung oder der Translationsmechanismen ist ein lohnendes Forschungsfeld um neue Ansatzpunkte für die Medikamentenentwicklung zu finden. Es ist offensichtlich, dass die wenigen bisher identifizierten RNA Bindeproteine in den apicomplexen Mitochondrien nicht die Vielzahl der Ereignisse in der Prozessierung, Stabilisierung und Translation der mitochondrialen RNAs erklären kann.

### **1.4 PPR- und OPR-Proteine sind dominante Faktoren in der organellären RNA Prozessierung und Stabilisierung in photoautotrophen Organismen**

#### **1.4.1 Struktur und Funktion von *Pentatricopeptide repeat* Proteinen (PPR)**

Die Genexpression der Organellen in Pflanzen ist stark abhängig von kernkodierten Proteinen. Eine in höheren Pflanzen stark verbreitete Familie sind die *Pentatricopeptide repeat* Proteine (PPR). In *Arabidopsis thaliana* sind etwa 450 PPR-Proteine beschrieben. Diese sind exklusiv im Kerngenom kodiert und werden post-translational in die Plastiden und Mitochondrien transportiert (Zusammengefasst in Barkan und Small 2014). Während in Pflanzen die PPR-Familie stark vergrößert ist, finden sich in anderen Organismen nur wenige PPR-Proteine. Im Menschen wurden lediglich 7 PPRs identifiziert (Lightowlers und Chrzanowska-Lightowlers 2013, zusammengefasst in Schmitz-Linneweber und Small 2008) und in *P. falciparum* nur zwei (Hicks et al. 2018). PPR-Proteine sind definiert durch ein degeneriertes, sich wiederholendes Motiv aus 35 Aminosäuren (AS). Dieses liegt bis zu dreißigmal in einem Protein vor und jedes Motiv formt zwei antiparallele Helices, während das gesamte Protein ein  $\alpha$ -Solenoid bildet. Die Struktur eines PPR-Proteins ist in Abbildung 4 gezeigt. Das Motiv der PPR-Proteine ist verwandt mit dem der TPR-Proteine (*tetratricopeptide repeat*) (Small und Peeters 2000, zusammengefasst in Barkan und Small 2014). Strukturell zeigen PPR Proteine eine Ähnlichkeit mit weiteren helikalen Proteinen, wie beispielsweise den PUF-Proteinen (*Pumilio and FBF homology*, zusammengefasst in Wang et al. 2018)



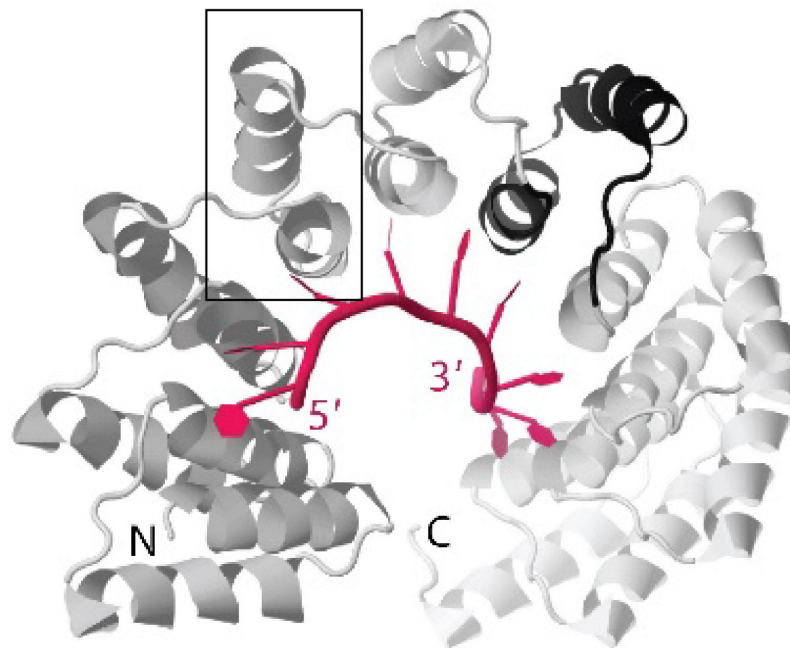


Abbildung 4: Struktur eines PPR-Proteins. Ein PPR-Motiv formt zwei antiparallel Helices (schwarze Box) und jedes Motiv erkennt eine Base in einem RNA-Molekül (RNA in rot). Das gesamte Protein liegt als  $\alpha$ -Solenoid vor (verändert nach Barkan und Small 2014).

und den DNA-bindenden TAL-Proteinen (*transcription activator-like*, zusammengefasst in Mak et al. 2013). Auch die in 1.4.2 beschriebenen OPR-Proteine gehören zu den PPR-ähnlichen Proteinen. Alle diese Proteine erkennen mit jedem Motiv des Proteins eine Base in einem DNA/RNA Molekül, wobei nur einige Aminosäuren des Motivs für die Erkennung der spezifischen Base verantwortlich sind (zusammengefasst in Hammani et al. 2014). Es konnte gezeigt werden, dass in PPR Motiven die AS 6 in der zweiten und AS 1 in der ersten Helix jedes PPR-Motivs entscheidend für die Basenerkennung sind (zusammengefasst in Barkan und Small 2014). Diese Positionen wurden genutzt, um aus PPR-Proteinen mit bekannter Bindestelle einen PPR-Bindungscode zu erstellen, der die Vorhersage von Bindestellen ermöglicht (Barkan et al. 2012). Die Verknüpfung dieses Codes mit der RNA Bindestelle wurde für mehrere PPR Proteine gezeigt, unter anderem für SOT1, welches in der rRNA Prozessierung von plastidären 23S – 4.5S rRNAs eine wichtige Rolle spielt (Wu et al. 2016).

Die genaue Funktion vieler PPR-Proteine in Pflanzen ist intensiv untersucht worden. Es konnte gezeigt werden, dass die PPR Proteine sehr spezifisch einzelsträngige RNAs binden, häufig in den nicht kodierenden Regionen (UTR, *untranslated regions*) von poly- und monocistronischen mRNAs. Durch diese Bindung werden RNA-Exonukleasen blockiert, wodurch die Stabilität der Transkripte erhöht wird (siehe 1.5, zusammengefasst in Schmitz-Linneweber und Small 2008). So konnte beispielsweise gezeigt werden, dass das Protein PPR10 in *Zea mays* im intergenischen Bereich zwischen den chloroplastidären Genen *atpH* und *psaJ-rpl33* bindet (Pfalz et al. 2009). Diese Bindung ist wichtig für die Definition und Stabilisierung der 5'- bzw. 3'-Enden der jeweiligen Transkripte in Zusammenarbeit mit den RNasen der Plastiden. Gleichzeitig führt die Bindung von PPR10 an die *atpH* mRNA zu einer Strukturänderung im Bereich der Ribosomenbindestelle, wodurch die Translation dieser mRNA gefördert wird (Prikryl et al. 2011). Ein ähnlicher Stabilisierungsmechanismus wurde für das PPR Protein HCF152 (High Chlorophyll Fluorescence 152) gezeigt. Hierbei werden die Transkriptenden zwischen *psbH-petB* stabilisiert (Zhelyazkova et al. 2012, Meierhoff et al. 2003). Die PPR-Proteinfamilie kann in zwei Untergruppen unterteilt werden. PPR10 und HCF152 gehören zu der P-Klasse der Proteine, die durch konstante 35 AS lange Motive definiert werden und neben dem Lokalisationspeptid am Amino-Terminus (N-Terminus) meist keine weiteren Domänen besitzen. Die andere Klasse der PPR-Proteine wird als PLS-Klasse bezeichnet, da die Länge der PPR-Motive variiert (L-Motiv 35-36 AS, S-Motiv 31 AS) (Lurin et al. 2004). Zusätzlich tragen die meisten PLS-Klasse Proteine zusätzliche Carboxy-terminale (C-Terminus) Domänen. Die meisten bisher untersuchten PLS-PPRs sind für die Editierung einzelner Transkripte verantwortlich (Kotera et al. 2005), allerdings wurden kürzlich auch PLS-PPR-Proteine beschrieben, die Transkripte stabilisieren (Ruwe et al. 2018, Hammani et al. 2016).

Auch außerhalb der photosynthetischen Organismen wurden PPR-Proteine studiert. Im Erreger der Afrikanischen Trypanosomiasis (Schlafkrankheit), *Trypanosoma brucei*, wurden 36 PPR-Proteine identifiziert. Davon besitzt der größte Teil eine mitochondriale Translokationssequenz. Interessant ist, dass in Präzipitationsexperimenten der mitochondriellen Ribosomen 21 PPR-Proteine mit den Ribosomen assoziiert waren (Mingler et al. 2006, Aphasizheva et al. 2011).

Zwei dieser ribosomalen PPR Proteine wurden tiefergehend charakterisiert (KRIPP1 und KRIPP8, *kinetoplast ribosomal PPR proteins*). Diese beiden Proteine binden in der kleinen Untereinheit des mitochondrialen Ribosoms und in *knock-down*-Experimenten von KRIPP1 konnte mit verringerter PPR-Proteinmenge eine Reduktion der *coxI* mRNA und des CO1 Proteins detektiert werden (Aphasizheva et al. 2015).

Die eingangs erwähnten humanen PPR Proteine sind alle mitochondrial lokalisiert und binden dort RNA. MRPS27 ist eines dieser PPRs und es wurde als Teil der kleinen Untereinheit der Ribosomen identifiziert. In *knock-down* Zellen von MRPS27 waren die RNA Level im Vergleich zur Kontrolle nicht verändert. Auf Proteinebene konnte eine reduzierte Translationseffizienz der mitochondrialen Proteine gezeigt werden (Davies et al. 2012). In der gleichen Studie wurde durch die Präzipitation des markierten MRPS27 festgestellt, dass neben der 12S rRNA auch die tRNA<sup>Glu</sup> mit dem PPR-Protein assoziiert. Ein Verlust von MRPS27 könnte die Rekrutierung dieser tRNA an den Ribosomen verlangsamen und so zu einer globalen Reduktion der mitochondrialen Translation führen (Davies et al. 2012).

### 1.4.2 Struktur und Funktion von *Octatricopeptide repeat* Proteinen (OPR)

Ähnlich zu den PPR-Proteinen der Pflanzen wurden in der einzelligen Grünalge *Chlamydomonas reinhardtii* helikale Proteine beschrieben, die strukturell den PPRs ähnlich sind, jedoch sind die degenerierten Motive mit 38–40 AS länger. Die Proteinfamilie wurden deshalb *Octatricopeptide-repeat* Proteine (OPR) genannt (Eberhard et al. 2011). In *C. reinhardtii* konnten bis zu 43 (Rahire et al. 2012) bzw. 42 Proteine (Eberhard et al. 2011) identifiziert werden, die multiple OPR-Motive besitzen. Die Lokalisationsvorhersage deutet, ähnlich wie bei PPR-Proteinen, auf eine Organellenlokalisierung hin. Die einzelnen Motive liegen hier ebenfalls in Wiederholungen vor und formen das von PPR-Proteinen bekannte Strukturmotiv aus zwei antiparallelen Helices. Obwohl strukturell ähnlich, unterscheiden sich OPR- und PPR-Motive auf Sequenzebene deutlich. In OPR-Motiven von *C. reinhardtii* gibt es konservierte Aminosäuren, die eine PPPEW-Domäne (Position 20-24) im OPR Motiv bilden (Rahire et al. 2012). Ob diese Domäne eine direkte Funktion hat ist unbekannt. Im Gegensatz zu PPR-Proteinen sind derzeit keine OPR-Proteine in Menschen identifiziert worden. Einige OPR-Proteine besitzen eine

C-terminale RAP-Domäne (ein Akronym für: *RNA-binding domain abundant in Apicomplexans*), eine RNA-Bindedomäne, die besonders häufig in den Apicomplexa gefunden wurde (Lee und Hong 2004). Besonders der C-terminale Teil der RAP-Domäne ist homolog zu den OPR-Motiven, sodass eine evolutionäre Verwandtschaft zwischen der RAP-Domäne und den OPR-Motiven vermutet wird (Eberhard et al. 2011, zusammengefasst in Hammani et al. 2014). In 16 der identifizierten OPR-Proteinen in *C. reinhardtii* konnte mindestens eine FAST-Domänen (*Fas-activated-Serine/Threonine kinase-like*) identifiziert werden (Eberhard et al. 2011, zusammengefasst in Hammani et al. 2014). Diese potentiellen Kinase-Domänen sind in Proteinen der FASTK-Proteinfamilie zu finden, die in den humanen Mitochondrien im RNA Metabolismus eine wichtige Rolle spielen (siehe 1.4.3, Jourdain et al. 2017). Derzeit ist nicht bekannt, ob die FAST-Domäne tatsächlich eine Kinaseaktivität hat, oder ob diese Domäne eine andere Funktion bzw. keine Funktion hat (Jourdain et al. 2017).

Während die in 1.4.1 beschriebenen PPR Proteine gut untersucht und charakterisiert sind, wurden bisher nur wenige OPR-Proteine charakterisiert. Bei der Untersuchung des OPR-Proteins Tab1 (für *translation of psaB mRNA*) in *C. reinhardtii* wurde durch eine Basendeletion in der kodierenden Sequenz des Proteins ein Translationsdefekt der plastidären *psaB*-mRNA identifiziert. Auch konnte in der gleichen Untersuchung gezeigt werden, dass Tab1 spezifisch die 5'-UTR Region der *psaB*-mRNA bindet (Rahire et al. 2012). Auch die Untersuchungen weiterer OPR-Proteine zeigten eine Funktion im RNA Metabolismus. Mutanten des OPR-Proteins Tbc2 (für *translation of psbC mRNA*) zeigten einen Translationsdefekt der chloroplastidären *psbC* mRNA (Auchincloss et al. 2002). Das Protein MCD1 (*maturation of the 3rd photosynthetic protein complex specifically affecting petD*) konnte ebenfalls mittels Deletionsmutanten charakterisiert werden. Hierbei wurde festgestellt, dass die Stabilität der *petD* mRNA stark vermindert wird, sodass 5'→3' Exonukleasen die RNA verdauen können (Murakami et al. 2005). In höheren Pflanzen ist mit dem Protein RAP (*RAP domain-containing protein*) bisher nur ein OPR-Protein beschrieben und charakterisiert worden. Dieses Protein beinhaltet 4 OPR-Motive und in einer funktionellen Studie zeigten Deletionsmutanten des Proteins einen Translationsdefekt der plastidären Transkripte. Die Mutanten haben einen extremen Phänotyp mit gebleichten Blättern und kümmerlichem Wachstum.

Neben den OPR-Motiven wurde am C-Terminus dieses Proteins eine RAP-Domäne identifiziert (Kleinknecht et al. 2014, Kleinknecht et al. 2016).

Zusammengefasst kann man sagen, dass ähnlich wie PPR Proteine auch OPR Proteine RNA bindende Proteine sind, die besonders in den Organellen von *C. reinhardtii* für die Translation und die RNA Stabilität spezifischer Transkripte verantwortlich sind. Ob OPR Proteine ähnlich wie in 1.5 beschriebenen cosRNAs generieren können, ist unbekannt.

### **1.4.3 Helikale Proteinfamilien spielen in den Organellen vieler Organismen eine wichtige Rolle**

Neben den oben beschriebenen OPR- und PPR-Proteinen gibt es weitere Proteine deren Struktur durch helikale, sich wiederholende Motive definiert werden. Zu diesen gehören die *tetratricopeptide repeat* (TPR) Proteine, welche strukturell ähnlich sind, jedoch eher für Protein-Protein Interaktionen als für Protein-RNA Interaktionen wichtig sind (Blatch und Lässle 1999). Nah verwandt zu den TPRs sind HAT-Proteine (HAT = *Half-a-Tetratricopeptide repeat*). Diese haben im Vergleich zu den TPRs einige konservierte Aminosäuren in den 34 AS langen Motiven, welche bei den TPR-Proteinen fehlen. HAT-Proteine liegen im Zellkern und im Zytosol vor, wo sie verschiedene Funktionen in der Reifung von RNAs übernehmen (zusammengefasst in Hammani et al. 2014). In humanen Mitochondrien ist neben den PPR-Proteinen eine weitere Proteingruppe für den RNA-Metabolismus wichtig. Diese besitzen zwei Kinase-ähnliche Domänen (FAST\_1 und FAST\_2) und werden in der FASTK-Proteinfamilie zusammengefasst (Jourdain et al. 2017). Die Domänen der FASTK-Proteine zeigen Ähnlichkeiten mit den OPR-Motiven (Eberhard et al. 2011). Die funktionelle Analyse der FASTK-Proteine deutet auf eine wichtige Rolle bei der Translation und der Prozessierung von Transkripten in humanen Mitochondrien hin (Boehm et al. 2017, Jourdain et al. 2015, Antonicka und Shoubridge 2015). Alle sechs Proteine dieser Familie besitzen, ähnlich wie einige OPR-Proteine, ein C-terminale RAP-Domäne (Jourdain et al. 2017). Diese wurde erstmals als potentielle RNA-Bindedomäne in Apicomplexa beschrieben (Lee und Hong 2004). Obwohl eine Verwandtschaft zwischen der RAP-Domäne und den OPR-Motiven vermutet wird (Eberhard et al. 2011), deutet die bioinformatische Strukturanalyse dieser Domäne auf eine D-(D/E)-X-K-

Nuklease-faltung hin (Boehm et al. 2017). Während der RAP-Domäne vermutlich eine katalytische Funktion zukommt, wurde vermutet, dass die spezifische RNA-Erkennung der Proteine durch Elemente im N-terminalen Bereich stattfindet (Jourdain et al. 2017). Dieser Proteinbereich ist dominiert von helikalen Elementen und es wurde postuliert, dass helikale Elemente ähnlich den OPR- bzw. PPR-Motiven eine Rolle spielen könnten (Jourdain et al. 2017, Hammani et al. 2014).

### **1.5 Kurze RNAs in den Organellen höherer Pflanzen können von PPR-Proteinen generiert werden**

Anhand des PPR-Proteins PPR10 aus Mais konnte erstmals ein Mechanismus für die Funktion von PPR-Proteinen als Stabilisator von mRNAs bestimmt werden. Dieses Modell nach Pfalz et al. 2009 ist in Abbildung 5 dargestellt. Nach diesem Modell binden PPR Proteine in den intergenischen Regionen von polycistronischen Transkripten. Durch die Prozessierung durch Endo- und Exonukleasen werden monocistronische Transkripte produziert, deren Enden durch PPR-Proteine stabilisiert werden. Ein PPR-Protein kann also sowohl die stromauf- als auch die stromabwärts der Bindestelle gelegenen RNAs stabilisieren. Die resultierenden mRNAs enden jeweils an der Bindestelle des PPR-Proteins. Dies führt zu überlappenden Transkriptenden des stromauf- und stromabwärts gelegenen Gens. Die PPRs fungieren bei der monocistronischen mRNA als Blocker, die das gebundene Transkriptende vor exonukleolytischem Abbau schützen und so die Halbwertszeit stark erhöhen. Wenn die mRNA durch bidirektionalen Verdau mit Exonukleasen abgebaut wird, bleibt die Bindestelle des PPR-Proteins weiter geschützt. Derartige kurze RNAs können mittels Hochdurchsatzsequenzierung transkriptomweit detektiert werden (Ruwe und Schmitz-Linneweber 2012). Die kurzen RNAs, die von potentielle Bindestellen der PPR-Protein abstammen, werden als cosRNAs (*clustered organellar short RNAs*) bezeichnet. Bei der Detektion der cosRNAs in der bioinformatischen Analyse müssen diese von anderen kurzen RNAs abgegrenzt werden. Die identifizierten cosRNAs zeichnen sich in der bioinformatischen Analyse meist durch ein scharfes 5'- und 3'-Ende aus. Zwischen den Enden befindet sich ein Plateau bei dem die Genomabdeckung durch die Sequenzierung konstant hoch ist. An den Enden der cosRNA kommt es zu einem

## Einleitung

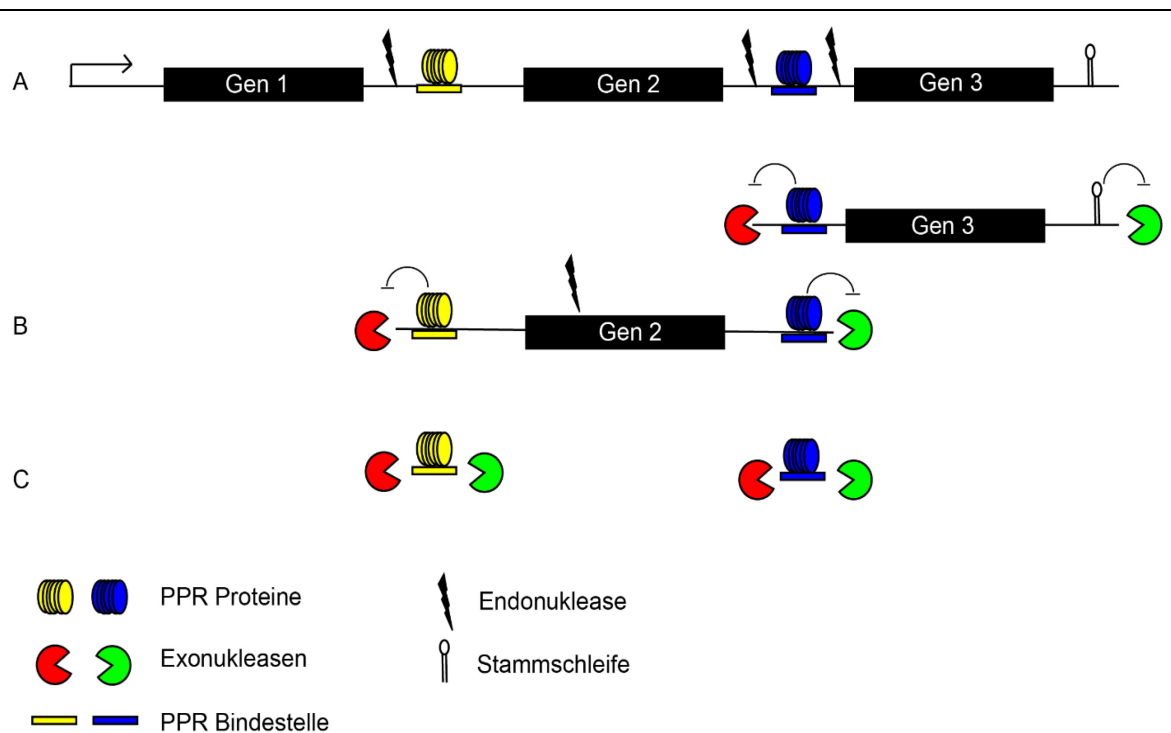


Abbildung 5: Modell für RNA Prozessierung in Plastiden: In den Plastiden höherer Pflanzen werden Gene als polycistronische Transkripte transkribiert. **A)** Eine polycistronische RNA wird in den intergenischen Regionen von PPR-Proteinen sequenzspezifisch gebunden. Die Prozessierung zu monocistronischen Transkripten erfolgt durch Endonukleasen. **B)** Durch Exonukleasen werden die UTRs der Transkripte sowohl in 5' → 3' als auch von 3' → 5' degradiert. Die codierende Sequenz wird durch die PPR-Proteine von diesem Abbau geschützt; auch Stammschleifen können diese Funktion ausführen. **C)** Nach dem vollständigen Abbau bleiben die PPR-Bindestellen als kleine RNA Fragmente erhalten, in HTS-Datensätzen können diese cosRNAs identifiziert werden. Modell nach Pfalz et al. 2009.

plötzlichen Abfall der Genomabdeckung. Diese Kriterien können genutzt werden um mit Programmen wie dem *sRNA-Miner* cosRNAs genomweit zu detektieren (Ruwe et al. 2016). Neben den PPR-Proteinen können auch stabile Sekundärstrukturen der RNAs, wie etwa Stammschleifen, eine stabilisierende Wirkung auf die RNA haben.

Auf Grundlage des oben beschriebenen Modells wurden frei verfügbare HTS-Datensätze analysiert und eine Vielzahl dieser cosRNAs konnten in den Chloroplasten von *A. thaliana* identifiziert werden (Ruwe und Schmitz-Linneweber 2012). Auch im Mitochondrium von Pflanzen und Algen wurden ähnliche RNA-Akkumulationen entdeckt (Ruwe et al. 2016). Diese Verknüpfung von cosRNAs mit RNA-bindenden Proteinen ermöglicht eine spannende und neuartige Methode, um

Interaktionspartner zu identifizieren. Durch die Analyse der RNA einer Nullmutante eines PPR-Proteins können also Rückschlüsse auf die Proteinbindestelle gezogen werden. In *A. thaliana* Deletionsmutanten des PLS-Klasse PPR-Proteins CRR2 (*Chlororespiratory Reduction 2*) wurde die RNA sequenziert und mit den Sequenzierungsdaten der RNA der Wildtyppflanze verglichen. Dabei wurden einigen cosRNAs identifiziert, die in der PPR-Deletionsmutante im Vergleich zum Wildtyp dramatisch reduziert vorlagen (Ruwe et al. 2018). In *Northern*-Hybridisierungen konnte das Fehlen dieser cosRNAs in der Mutante bestätigt werden, während andere cosRNAs keine Veränderung zeigten (Ruwe et al. 2018). In anderen PPR-Mutanten wurden die cosRNAs genutzt, um bei bekannter Ziel-RNA des PPR-Proteins die genaue Bindestelle zu identifizieren. So wurde für das PPR-Protein MTSF2 in *A. thaliana* die genaue Bindestelle im *nad1* Transkript bestimmt, indem die Abwesenheit einer cosRNAs in dieser Mutante gezeigt wurde (Wang et al. 2017).

### 1.6 Zielstellung

Über den RNA-Metabolismus und die Regulation der Expression in den Organellen von *Plasmodium* ist wenig bekannt. Aufgrund der endosymbiontischen Abstammung des Apicoplasten und des Mitochondriums soll in dieser Arbeit die potentielle Rolle von helikalen Proteinen wie OPR- und PPR-Proteinen untersucht werden. Diese Proteine übernehmen in den Organellen von höheren Pflanzen und Algen eine wichtige Rolle im RNA-Metabolismus. Eine Möglichkeit um potentielle Bindestelle dieser Proteine zu identifizieren, ist die Analyse von RNA-Sequenzierungsdaten. In höheren Pflanzen konnte an den potentiellen Proteinbindestellen die Akkumulation einer kurzen RNA detektiert werden. In dieser Arbeit soll hierzu die zelluläre RNA von *P. falciparum* durch Hochdurchsatzsequenzierungen analysiert werden. In der bioinformatischen Auswertung sollen besonders die Organellen von *P. falciparum* auf Akkumulationen von kurzen RNAs untersucht werden, die den cosRNAs in den Organellen von Pflanzen ähnlich sind. Einige ausgewählte, potentielle cosRNAs sollen durch komplementäre Methoden bestätigt werden.



In *Plasmodium* wurden zwei PPR-Proteine identifiziert (siehe 1.4.1). Andere RNA-bindende Proteine, die am RNA-Metabolismus der Organellen beteiligt sein könnten, sind derzeit nicht bekannt. Aufgrund des endosymbiontischen Ursprungs der Organellen soll in einer genomweiten Suche nach Proteinen mit helikalen repetitiven Elementen gesucht werden, wie sie in photoautotrophen Organismen in großer Zahl gefunden werden. Der Fokus liegt hier besonders auf OPR-ähnlichen Proteinen. Gefundene Proteine sollen bioinformatisch detaillierter untersucht werden. Ausgesuchte Proteine sollen anschließend biochemisch *in vitro* und genetisch *in vivo* funktionell untersucht werden, um ein tieferes Verständnis der Funktion dieser Proteine im RNA-Metabolismus der Organellen der Parasiten zu erhalten.

## 2. Material und Methoden

### 2.1 Materialien

Eine Liste der Hersteller und Anbieter ist im Anhang (Anhang Tabelle 1) aufgeführt.

#### 2.1.1 Parasiten

Für die Sequenzierung von kurzen RNAs, sowie der Untersuchung der mitochondrialen rRNA-Fragmente wurde *Plasmodium falciparum* NF54 verwendet. Die Parasiten wurden in der Vektorbiologie Gruppe unter der Leitung von Dr. Elena A. Levashina (Max-Planck-Institut für Infektionsbiologie, Berlin) kultiviert und großzügig zur Verfügung gestellt. Für die Deletionsstudien, sowie für die mikroskopische Untersuchung der zellulären Lokalisation von Proteinen, wurden Parasiten des Stamms *Plasmodium berghei* ANKA verwendet. Diese Arbeiten wurden in der Molekularen Parasitologie unter der Leitung von Prof. Dr. Kai Matuschewski (Humboldt Universität zu Berlin) durchgeführt.

#### 2.1.2 Mäuse

In den Experimenten mit *Plasmodium berghei* wurden *Mus musculus* NMRI (Naval Medical Research Institute, Charles River Laboratories, Deutschland) verwendet. Die Experimente wurden in Übereinstimmung mit dem Deutschen Tierschutzgesetz durchgeführt.

#### 2.1.3 Chemikalien

Soweit nicht anders vermerkt, wurden alle Lösungen und Puffer mit destilliertem und deionisiertem Wasser hergestellt (A. bidest. PURE-Lab-Ultra-Anlage, Veolia). Alle Chemikalien, wenn nicht anders angegeben, wurden von Carl Roth, Sigma-Aldrich, Merck und Roche bezogen. Für die Detektion von RNAs über spezifische Oligonukleotid-Sonden wurde radioaktives Material von Hartmann Analytic GmbH bezogen.

#### 2.1.4 Enzyme

Alle verwendeten Enzyme, soweit nicht anders angegeben, wurden von Thermo Fisher, Qiagen und New England Biolabs (NEB) bezogen.

#### 2.1.5 Bakterienstämme

Rekombinante Plasmide wurden in *Escherichia coli* TOP10 (Invitrogen) und XL-1 blue (Agilent) propagiert. Für die heterologe Expression von rekombinanten Proteinen wurde *E. coli* BL21-CodonPlus(DE3)-RIPL (Agilent) genutzt.

#### 2.1.6 Plasmide

In der folgenden Tabelle sind die verwendeten Plasmide, deren Anwendung und die Herkunft/Bezugsquelle aufgelistet.

Tabelle 1: Liste der verwendeten Plasmide. Neben den Namen ist auch die Anwendung und die Herkunft des jeweiligen Plasmids angegeben.

Plasmid	Anwendung	Herkunft
pMAL-TEV	Rekombinante Expression von Proteinen in <i>E. coli</i>	NEB, modifiziert von Prof. Dr. A. Barkan, USA
pBAT1-G	Markierung von Proteinen für Lokalisationsstudien in <i>P. berghei</i>	AG Molekulare Parasitologie, HU-Berlin
pBAT-M6	Deletionsvektor für Proteine in <i>P. berghei</i>	AG Molekulare Parasitologie, HU-Berlin

#### 2.1.7 Antikörper

Für die Immundetektion von rekombinanten Fusionsproteinen wurde ein monoklonaler Antikörper (*Anti-MBP Monoclonal Antibody*, NEB) verwendet. Der Antikörper bindet spezifisch an ein Antigen des MBP-Proteintag. In der Immundetektion wurde der Antikörper 1:10.000 verdünnt (1x TBST und 2% Trockenmilchpulver) eingesetzt. Die Detektion erfolgte über einen, mit einer Peroxidase (HRP = *Horseradish peroxidase*) konjugierten, sekundären Antikörper (*Mouse-HRP*, GE-Healthcare). Dieser erkennt spezifisch das Maus-IgG des Anti-MBP-Antikörpers und wurde ebenfalls 1:10.000 verdünnt (1x TBST).

### **2.1.8 Oligonukleotide**

Alle verwendeten Oligonukleotide wurden kommerziell synthetisiert (Invitrogen oder Sigma-Aldrich). Eine vollständige Liste der verwendeten Oligonukleotide ist im Anhang hinterlegt (siehe Anhang Tabelle 2).

### **2.1.9 Software**

Die für Dokumentation, Auswertung, Visualisierung und Experimentplanung verwendeten Computerprogramme und Software sind im Anhang (siehe Anhang Tabelle 3) aufgelistet.

### **2.1.10 Datenbanken**

Für die allgemeinen Protein- und Genominformationen wurde die PlasmODB<sup>1</sup> verwendet (Aurrecochea et al. 2009). Für die Proteome der Motivsuchen wurde die Uniprot<sup>2</sup> Datenbank genutzt. Zusätzliche Informationen wurden über NCBI<sup>3</sup> bezogen. Für die Suche nach konservierten Proteindomänen wurde die PFAM-Datenbank<sup>4</sup> genutzt.

## **2.2 Methoden**

### **2.2.1 Inaktivierung von biologischem Material**

Die Inaktivierung von biologischem Material (besonders GVOs, gentechnisch veränderte Organismen) und die Sterilisation von Lösungen und Medien erfolgte durch Dampfsterilisation. Die Sterilisation erfolgte bei 120°C und 55 kPa für 20 Minuten im Systec VE 75 Autoklaven (Systec).

---

<sup>1</sup> <http://www.plasmodb.org>

<sup>2</sup> <https://www.uniprot.org/>

<sup>3</sup> <https://www.ncbi.nlm.nih.gov/>

<sup>4</sup> <https://pfam.xfam.org/>

## **2.2.2 Kultivierung von *Escherichia coli***

### **2.2.2.1 Kultivierung auf Agar-Agar-LB Platten**

Für die Antibiotika-Selektion von rekombinanten Plasmiden in *E. coli* Stämmen, wurden die Bakterien auf LB-Agar Platten mit Selektionsmarkern ausplattiert. Die Inkubation erfolgte für mindestens 16 Stunden bei 37°C. Für alle hier verwendeten Plasmide wurde als Selektionsmarker Carbenicillin (100 µg/ml Konzentration) verwendet. Das Antibiotikum wurde dem LB-Agar unmittelbar vor dem Gießen der Platten zugesetzt.

LB-Agar-Platten:                    1 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 1 % (w/v) NaCl, pH 7,0 mit NaOH, 1,5 % (w/v) Agar-Agar

### **2.2.2.2 Kultivierung in flüssigem LB-medium**

Die heterologe Herstellung von rekombinanten Proteinen und die Produktion der Deletions- bzw. Markierungskonstrukte erfolgte mit *E. coli* Zellen in flüssigem LB-Medium. Nach der Transformation der Bakterien mit der Plasmid-DNA wurden die Zellen auf LB-Selektionsplatten über Nacht kultiviert. Eine einzelne Kolonie wurde steril in flüssiges LB-Medium (mit Selektionsmarker Carbenicillin 100 µg/ml) gegeben und unter starkem Schwenken über Nacht kultiviert (37°C und 220 rpm (*rounds per minute*)). Die Zellen wurden mindestens 16 Stunden inkubiert. Die Ernte der Zellen erfolgte durch Zentrifugation (5.000x g für 5 min).

LB-Medium:                        1 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 1 % (w/v) NaCl, pH 7,0 mit NaOH

## **2.2.3 Kultivierung von *Plasmodium falciparum***

Die Kultivierung von *P. falciparum* NF54 wurde in der Arbeitsgruppe Vektorbiologie von Dr. Elena A. Levashina am Max-Planck-Institut für Infektionsbiologie (Berlin) nach publizierten Protokollen durchgeführt (Trager und Jensen 1976). Die Kultivierung erfolgte in Dauerkulturen. Die Kulturen wurden alle

2 bzw. 3 Tage geteilt und in frisches Wachstumsmedium (RPMI 1640, Invitrogen) transferiert.

#### 2.2.4 Amplifikation von DNA-Molekülen mittels PCR

Die Amplifikation von DNA Fragmenten erfolgte durch Polymerase-Kettenreaktion (PCR) mit thermostabilen Polymerasen. Da sowohl *P. berghei* als auch *P. falciparum* im Kerngenom nur einen GC-Gehalt von etwa 20 % haben, wurde die Elongationstemperatur auf 65°C reduziert (Su et al. 1996). Die Elongationszeit wurde dafür von 60 s auf 90 s pro 1000 bp des erwarteten Amplifikats erhöht. DNA Fragmente, die für Klonierungen bestimmt waren, wurden mit einer Q5® High-Fidelity DNA Polymerase (NEB) amplifiziert (Programm und Reaktion nach Herstellerangaben). Alle anderen PCRs wurden mit *Taq*-Polymerase (Eigenherstellung nach Desai und Pfaffle 1995) durchgeführt. Als Standardreaktion wurde 1x PCR Puffer, 0,2 mM dNTPs (Thermo Fisher), je 0,2 µM des *forward* und *reverse* Primers, sowie 1:40 verdünnte *Taq*-Polymerase angesetzt. Die Konzentration der DNA-Vorlage wurde mittels Spektroskopie bestimmt (siehe 2.2.8) und die entsprechende Verdünnung in die Reaktion eingesetzt (für Plasmide maximal 5 ng, für gDNA maximal 100 ng). Die Schmelztemperatur für die Primer wurde mit „*NEB tm calculator*<sup>5</sup>“ berechnet. Ein Beispiel-PCR-Programm ist in der folgenden Tabelle aufgeführt. Je nach benötigter Produktmenge wurde die Anzahl der Zyklen zwischen 25-40 variiert (ein Zyklus: Denaturierung, Hybridisierung, Elongation).

Schritt	Temperatur	Dauer
Initiale Denaturierung	95°C	3 Min
Denaturierung	95°C	30 Sek
Primer-Hybridisierung	variabel	30 Sek
Elongation	65°C	75 Sek pro kb
Finale Elongation	65°C	7 Min
Lagerung	8°C	∞

Zur Qualitätsprüfung wurde eine Probe der Reaktion mit 1/10 Vol. 10x DNA-Ladepuffer versetzt und mittels Gelelektrophorese analysiert (siehe 2.2.7).

<sup>5</sup> <https://tmcalculator.neb.com>

10x PCR-Puffer:	200mM Tris-HCl pH 8.8, 100mM KCl, 100mM (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , 20mM MgSO <sub>4</sub> , 1 % Triton X-100
10x DNA Ladepuffer:	25 % Ficoll (400), 0,42 % (w/v) Bromphenolblau, 0,42 % (w/v) Xylencyanol

### 2.2.5 Transformation von chemisch kompetenten *Escherichia coli*

Von einer Ligationsreaktion wurden 2 – 5 µl eingesetzt um 50 µl chemisch-kompetente *E. coli* Bakterien zu transformieren. Die kompetenten Zellen wurden auf Eis langsam aufgetaut. Nach dem Auftauen wurde die DNA zugegeben und für 45 min auf Eis inkubiert. Um die Aufnahme der DNA zu fördern, wurden die Bakterien für 45–60 s einem Hitzeschock bei 42°C im Heizblock ausgesetzt. Nach kurzer Inkubation auf Eis wurden die Zellen mit 800 µl SOC-Medium (Hanahan 1983) für 1 Stunde inkubiert (37°C und schütteln bei 220 rpm). Anschließend wurden unterschiedliche Mengen auf LB-Agarplatten (modifiziert nach Bertani 1951) (Selektionsmarker 100 µg/ml Carbenicillin) ausplattiert und über Nacht inkubiert.

SOC-Medium:	2 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 10 mM NaCl, 2,5 mM KCl, 10 mM MgSO <sub>4</sub> , pH 7,0 mit NaOH; 10 mM MgCl <sub>2</sub> und 20 mM Glukose
LB-Agar-Platten:	1 % (w/v) Trypton, 0,5 % (w/v) Hefeextrakt, 1 % (w/v) NaCl, pH 7,0 mit NaOH; 1,5 % (w/v) Agar-Agar

### 2.2.6 Kolonie-PCR

Die Überprüfung von inserierter DNA in Plasmiden wurden durch eine Kolonie-PCR überprüft. Hierzu wurden Einzelkolonien mit steriler Pipettenspitze aufgenommen und in 20 µl *A. bidest.* resuspendiert. Die Zellsuspension wurde als DNA-Vorlage in einer PCR-Reaktion eingesetzt (siehe 2.2.4). Als Primer in der Kolonie-PCR wurden DNA-Oligonukleotide eingesetzt, die komplementär zur Sequenz des Plasmids stromauf bzw. stromab der Insertionsstelle waren. Das PCR-Programm entsprach dem in 2.2.4 beschriebenen. Für die Überprüfung der Reaktion wurden 8 µl jeder Reaktion mit 1/10 Vol. 10x DNA-Ladepuffer vermischt und in einem 1%igen Agarosegel aufgetrennt (siehe 2.2.7). Parallel dazu wurde ein

Größenstandard auf demselben Gel aufgetragen (GeneRuler 1kb, GeneRuler 50bp, Thermo Fisher). Die Dokumentation erfolgte unter UV-Licht (siehe 2.2.7). Kolonien, die das gewünschte Fragment zeigten, wurden in 6 ml LB-Medium mit passendem Selektionsmarker angeimpft und über Nacht kultiviert. Anschließend wurden die Plasmide isoliert und sequenziert (siehe 2.2.9).

## 2.2.7 Gelelektrophorese von Nukleinsäuren

Die Längenanalyse, Qualitätsbewertung und Aufreinigung von Nukleinsäuren wurde mittels einer nicht-denaturierenden Agarose-Gelelektrophorese durchgeführt. Dazu wurde Agarose (Biozym) zu 1-2% (w/v) in 1x TAE Puffer aufgekocht und mit 0,2 µg/ml Ethidiumbromid versetzt. Die heiße Agaroselösung wurde in horizontale Gelkammern (PeqLab, PerfectBlue-Gelsystem Mini bzw. -Midi) gegossen und ein Probenkamm eingesetzt. Die DNA-Proben wurden mit 1 Vol. 10x DNA-Ladepuffer vermischt. RNA-Lösungen wurden mit 1 Vol. RNA-Ladepuffer versetzt. Die RNA-Proben wurden anschließend durch Hitze denaturiert (75°C für 10 min im Heizblock) und bis zum Auftragen auf das Gel auf Eis gelagert. Zu den jeweiligen Proben wurde parallel ein Längenstandard (GeneRuler 1kb oder GeneRuler 50bp, Thermo Fisher) auf das Gel geladen. Der Gellauf erfolgte in horizontalen Gelkammern in 1x TAE Puffer und mit einer Spannung von 5 – 10 V/cm Gellänge (Sambrook und Russell 2001). Zur Detektion wurden die Nukleinsäuren durch UV-Licht sichtbar gemacht und dokumentiert (Gel Doc™ XR, Bio-Rad).

1x TAE Puffer:	40 mM Tris, 20 mM Acetat, 1 mM EDTA
10 x DNA-Ladepuffer:	25 % Ficoll (400), 0,42 % (w/v) Bromphenolblau, 0,42 % (w/v) Xylencyanol
RNA-Ladepuffer:	95 % deionisiertes Formamid, 1 mM EDTA, 0,02 % (w/v) SDS, Spuren von Bromphenolblau und Xylencyanol

## 2.2.8 Spektroskopische Analyse von Nukleinsäuren

Für die Mengenbestimmung und die Qualitätskontrolle von Nukleinsäuren wurde eine spektroskopische Analyse mittels NanoDrop™ One (Thermo Fisher)



durchgeführt. Die Qualität und Integrität von RNA wurde zusätzlich mit Gelelektrophorese bestimmt.

### **2.2.9 Isolation und Aufreinigung von PCR Produkten/Plasmiden**

Um DNA Moleküle aus PCR Reaktionen und Plasmid-DNA aus Bakterien zu isolieren, wurden die GeneJET PCR Purification bzw. GeneJET Plasmid Miniprep Kits (Thermo Fisher) nach Herstellerangaben verwendet. Die gereinigte DNA wurde spektroskopisch analysiert und in Folgereaktionen verwendet.

### **2.2.10 Verwendung von T4-DNA Ligase**

Für die Ligation von zwei DNA Fragmenten während der Klonierung wurde eine T4-DNA-Ligase (NEB) verwendet, welche die Verknüpfung von 3'-hydroxyl-Gruppen mit 5'-Phosphat-Gruppen katalysiert. Die Reaktion erfolgte nach Herstellerangaben über Nacht. Das Reaktionsvolumen betrug 20 µl und nach abgeschlossener Ligation wurde die Ligase durch Erhitzen inaktiviert (20 min, 65°C). Die Ligationsreaktion wurde direkt für die Transformation von chemisch kompetenten *E. coli* TOP10 bzw XL-1 blue verwendet (siehe 2.2.5).

### **2.2.11 Verwendung von Restriktionsendonukleasen**

Die hier verwendeten Restriktionsendonukleasen wurden entweder von Thermo Fisher oder NEB bezogen und nach Herstellerangaben verwendet. Für die Kombination von Enzymen wurden die von den Herstellern bereitgestellten Programme verwendet (NEB: *NEBcloner*, Thermo Fisher: *DoubleDigest Calculator*). Die verwendeten Endonukleasen sind in den entsprechenden Methoden aufgeführt (siehe 2.2.29, 2.2.34, 2.2.35).

### **2.2.12 Extraktion von RNA aus organischem Material**

Die Extraktion von RNA aus Zellen und aus wässrigen Lösungen erfolgte durch Direct-zol™-RNA Mini-PrepPlus-Kit (Zymo Research). Nach der Aufreinigung und Lyse der Blutzellen (siehe 2.2.32) wurden die zentrifugierten Parasiten in 100 µl 1x PBS aufgenommen und mit 4 Vol. YelZol (basierend auf Chomczynski und Sacchi 2006) versetzt. Zur Unterstützung der Zelllyse wurde die Lösung 15 – 20 mal durch eine Spritze mit aufgesetzter Kanüle (0.45 x 12 mm, Braun) passiert. Für die

Extraktionen von RNA aus wässrigen Lösungen (z.B. *in vitro* Transkriptionsreaktionen) wurden die Lösungen mit 4 Vol. YelZol vermisch. Die Isolation der RNA erfolgte nach Herstellerangaben des Direct-zol™-RNA Mini-PrepPlus-Kits (Zymo Research).

YelZol: 4 M Guanidiniumthiocyanat, 25 mM Tris-HCl (pH 4,5), 0,5 % (w/v) Sarkosyl, 100 mM 2-Mercaptoethanol, 1,5 mM Hydroxyquinolin, 2 % (v/v) Isoamylalkohol, 5 mM EDTA, 50 % (v/v) Aqua-Phenol

### **2.2.13 Extraktion von Nukleinsäuren mit Phenol/Chloroform**

Die Extraktion von DNA/RNA aus wässrigen Lösungen wurde mittels Extraktion mit Phenol/Chloroform durchgeführt (Sambrook und Russell 2001). Die Lösung wurde hierfür mit 1 Vol. Phenol/Chloroform/Isoamylalkohol (Verhältnis 25/24/1, pH 7,5 – 8, Roth) auf dem Wirbelmischer vermischt. Zur Phasentrennung wurden die Proben für 5 min bei maximaler Geschwindigkeit in der Tischzentrifuge zentrifugiert. Die obere wässrige Phase wurde in ein neues Reaktionsgefäß transferiert und mit 1 Vol. Chloroform/Isoamylalkohol (Verhältnis 24/1) vermischt. Die Phasentrennung erfolgte wie oben beschrieben. Die wässrige Phase wurde in ein neues Reaktionsgefäß transferiert und wie in 2.2.14 beschrieben präzipitiert.

### **2.2.14 Präzipitation von Nukleinsäuren mit Ethanol**

Zum Konzentrieren und Reinigen von Nukleinsäuren wurde eine Präzipitation mit Ethanol durchgeführt. Die Nukleinsäurelösung wurde mit 1/10 Vol. 3 M NaAcetat und 3 Vol. 96% Ethanol versetzt und gründlich vermischt. Zur Unterstützung der Fällung wurden außerdem 1-2 µl GlycoBlue (Ambion) als Fällungshelfer zugesetzt. Der Ansatz wurde für mindestens 1 Stunde bei -20°C inkubiert. Die Lösung wurde anschließend in der Tischzentrifuge bei 21.000 x g und 4°C für 20 Min. zentrifugiert. Der Überstand wurde verworfen. Um Salze zu entfernen, wurde der Niederschlag zweimal mit kaltem 75 % Ethanol gewaschen. Das Ethanol des letzten Waschschritts wurde vollständig abgenommen und die präzipitierten Nukleinsäuren für 5-10 min an der Luft getrocknet. Abschließend wurden die Nukleinsäuren in gefiltertem *A. bidest.* gelöst. Zur Qualitätskontrolle wurden die Nukleinsäuren in

einer Gelelektrophorese analysiert (siehe 2.2.7). Die Mengen wurden durch spektroskopische Analyse bestimmt (siehe 2.2.8).

### **2.2.15 *In vitro* Transkription**

Für die Herstellung von Transkripten für RNA-Bindestudien wurde eine in T7-DNA abhängige RNA Polymerase (NEB) genutzt. Das mitochondriale Genom von *P. berghei* wurde dazu in 4 Fragmente unterschiedlicher Größe eingeteilt und mittels PCR amplifiziert. Die amplifizierten Fragmente hatten eine Länge von 1977 bp, 1704 bp, 1451 bp und 1200 bp (siehe auch Abbildung 20). Die verwendeten Oligonukleotide sind in Anhang Tabelle 2 aufgeführt. Gleichzeitig wurde über die verwendeten Oligonukleotide ein T7-Promotor (<sup>5'</sup>TAATACGACTCACTATAG<sup>3'</sup>) auf dem jeweiligen Strang eingefügt. Die PCR Produkte wurden gereinigt (siehe 2.2.9). Die *in vitro* Transkription erfolgte nach Herstellerempfehlungen über Nacht bei 37°C (T7-Polymerase, NEB). Zum Verdau der DNA wurden der Reaktion 4 U TURBO™ DNase (Invitrogen) zugefügt und für 30 min bei 37°C inkubiert. Die RNA wurde anschließend wie in 2.2.12 beschrieben aus der Reaktion extrahiert. Zur Überprüfung wurde die RNA spektroskopisch vermessen (2.2.8) und in einem denaturierenden Agarosegel analysiert (2.2.16).

### **2.2.16 Denaturierende RNA-Gelelektrophorese**

Zur Analyse von RNA-Lösungen und zur Vorbereitung des Transfers auf Membranen wurde eine denaturierende RNA-Gelelektrophorese durchgeführt. Hierzu wurden die RNA-Lösungen mit 2,5 - 3 Vol. RNA-Probenpuffer gemischt und 1/10 Vol. 10x Ladepuffer zugefügt. Die Proben wurden anschließend für 15 min bei 75°C denaturiert und bis zum Gellauf auf Eis gelagert. Ein passender Größenstandard wurde gleichbehandelt (RiboRuler High Range RNA Ladder bzw. RiboRuler Low Range RNA Ladder, Thermo Fisher). Die Gellösung wurde frisch angesetzt und das Gel gegossen. Der Gellauf erfolgte in horizontalen Elektrophoresekammern (Midi Large, neoLab) in einem Eisbad mit 1x MOPS Puffer (pH 7,0) bei einer Spannung von 5 – 7 V/cm. Zur Dokumentation wurde das Gel unter UV-Licht im Gel Doc™ XR (Bio-Rad) fotografiert. Anschließend wurde nach Bedarf ein Transfer auf eine Membran zur spezifischen Detektion bestimmter RNAs durchgeführt (siehe 2.2.17).

10x MOPS:	200 mM MOPS, 50 mM Natriumacetat, 10 mM EDTA, pH 7 bzw. 8 mit NaOH.
Gellösung:	1,2 g Agarose (Certified™ <i>Molecular Biology Agarose</i> , Bio-Rad) in 90 ml <i>A. dest</i> aufgekocht, auf 65°C abkühlen, 12 ml 10x MOPS Puffer (pH 8,0) und 20 ml 37 % Formaldehydlösung zufügen.
RNA-Probenpuffer:	65% (v/v) deionisiertes Formamid, 22 % (v/v) Formaldehyd-lösung (37 %), 13 % (v/v) 10X MOPS Puffer (pH 8).
10x Ladepuffer:	50 % (v/v) Glycerol, 10 mM EDTA, 0,25 % (w/v) Bromphenolblau, 0,25 % (w/v) Xylencyanol, 0,5 µg/µl Ethidiumbromid.

### 2.2.17 Transfer von RNA auf eine Nylonmembran (*Northern-blot*).

Nach der elektrophoretischen Auftrennung von RNA-Lösungen in denaturierenden Agarosegelen wurden die RNAs mittels passiver Kapillarkräfte auf eine Nylonmembran (Hybond-N Membran, GE-Healthcare) transferiert (Sambrook und Russell 2001). Als Transferpuffer wurde 5x SSC verwendet. Nach abgeschlossenem Transfer wurde die transferierte RNA durch UV-Bestrahlung auf der Membran verankert (250 mJ/cm<sup>2</sup>, GS-GeneLinker-UV-Kammer, Bio-Rad). Zur Verifizierung des Transfers und als Kontrolle der Ladung, wurde die Membran in Methylenblaulösung für 5 min gefärbt und mehrmals mit *A. dest.* bis zur gewünschten Intensität entfärbt. Anschließend wurde die Membran in der Hybridisierung mit radioaktiv markierten Sonden verwendet oder trocken bei RT gelagert.

5x SSC:	0,75 M NaCl, 75 mM Natriumcitrat, pH 7,0 mit HCl
Methylenblaulösung:	0,03 % (w/v) Methylenblau, 300 mM NaAc (pH 5,2)

## **2.2.18 Markierung, Hybridisierung und Detektion von radioaktiven Sonden**

### **2.2.18.1 Radioaktive Markierung von DNA-Oligonukleotiden**

Für den Nachweis spezifischer RNA-Fragmente wurden DNA-Oligonukleotide durch Markierung mit  $^{32}\text{P}$  am 5'-Ende markiert. Hierzu wurde eine Polynukleotidkinase (PNK, Thermo Fisher) nach Herstellerangaben verwendet. Als Substrat für die Markierung diente 50  $\mu\text{Ci}$  [ $\gamma$ - $^{32}\text{P}$ ]-ATP (Hartmann Analytics) mit dem ein entsprechendes Oligonukleotid markiert wurde (Anhang Tabelle 2). Die Reaktion wurde über Gelfiltrationssäulen aufgereinigt (illustra Microspin G-25 Säulen, GE-Healthcare).

### **2.2.18.2 Hybridisierung und Detektion der Sonden**

Membranen mit RNA wurden in Church-Puffer für 1 Stunde bei 37°C prähybridisiert. Die radioaktiv markierte Sonde wurde für 5 min bei 95°C denaturiert und hinzugefügt und über Nacht bei 37°C mit der Membran hybridisiert. Die Membran wurde zweimal für 10 min bei 37°C mit 1x SSC und 0,1 % (w/v) SDS gewaschen. Die Detektion der Sonden erfolgte über ein Phosphoimagersystem (PMI FX, Eraser Screen-K, Quantity-One-Software, Bio-Rad).

Church-Puffer:                      0,5 M Natriumphosphat Puffer (pH 7,0), 7 % (w/v) SDS,  
   1 mM EDTA

## **2.2.19 Denaturierende RNA Urea-Polyacrylamid-Gelelektrophorese (Urea-PAGE)**

Die Auftrennung von RNA in Polyacrylamidgelen ermöglicht eine hohe Auflösung im Bereich von kürzeren RNAs (zwischen 20 ~ 200 nt). Hierfür wurde ein PAGE-Gel mit Urea als denaturierende Substanz verwendet. Die RNA-Lösungen wurden mit 1 Vol. RNA-Ladepuffer vermischt und für 15 min bei 75°C denaturiert. Ein Größenstandard wurde parallel gleichbehandelt. Als Größenstandard wurden DNA-Oligonukleotide mit 22, 33 und 44 nt Länge verwendet (DNA wandert etwa 10 % schneller als RNA, Sambrook und Russell 2001). Zur Einordnung größerer RNAs wurde RiboRuler Low Range RNA Ladder (Thermo Fisher) genutzt. Der Gellauf

erfolgte bei 25 V/cm in 1x TBE-Puffer. Der Gellauf wurde gestoppt, als die dunkelblaue Bromphenolblau-Bande (im RNA-Ladepuffer enthalten) bis ca. 1 cm vor das Gelende gewandert war. Anschließend wurde das Gel für 10 min in 1 x TBE mit ~0,2 µg/ml Ethidiumbromid inkubiert und einmal mit 1x TBE gewaschen. Die Dokumentation erfolgte unter UV-Licht im Gel Doc™ XR (Bio-Rad).

Gellösung:	1x TBE, 8 M Urea (MP Biochemicals), 15 % Acrylamid (29:1 Acrylamid:Bisacrylamid), 0,5 % (v/v) TEMED, 0,05 % (w/v) APS
1x TBE:	89 mM Tris, 89 mM Borsäure, 20 mM EDTA, pH 8,0
RNA-Ladepuffer:	95 % deionisiertes Formamid, 1 mM EDTA, 0,02 % (w/v) SDS, Spuren von Bromphenolblau und Xylencyanol
Urea-PAGE-Lösung:	0,1 Vol. 10x TBE, 8 M Urea (MP Biochemicals), 15 % Polyacrylamid, 0,5 % (v/v) TEMED, 0,05 % (w/v) APS.

### 2.2.20 Transfer von kurzen RNAs aus der Urea-PAGE auf Nylonmembranen

Um spezifische kurze RNA-Fragmente zu detektieren, wurden die in der Urea-PAGE aufgetrennten RNA-Proben mittels Tank-*blot* (Mini Trans-Blot Cell, Bio-Rad) elektrophoretisch auf eine Nylonmembran transferiert. Der Transfer wurde in 0,5 x TBE Puffer und für 1 Stunde bei 80 V durchgeführt. Um die kurzen RNAs fest auf der Membran zu verankern (*crosslink*), wurde ein chemischer *crosslink* durchgeführt. Als *crosslinking*-Substanz wurde dazu 1-Ethyl-3-(3-Dimethylaminopropyl)Carbodiimid-Hydrochlorid (EDC) verwendet (Pall und Hamilton 2008). Durch diesen *crosslink* wird die Detektion von kurzen RNAs <40 nt im Vergleich zum Verankern mit UV-Licht dramatisch verbessert (Pall et al. 2007). Ein Filterpapier wurde mit *Crosslinking*-Lösung getränkt und die Membran mit der RNA abgewandten Seite darauf platziert. In Cellophan eingeschlagen, wurde der *crosslink* für 1 Stunde bei 60°C durchgeführt. Überschüssige *crosslinking*-Lösung wurde mit *A. dest.* abgespült. Zur Überprüfung des Transfers und als Ladekontrolle wurde die RNA in einer Methylenblaulösung (2-5 min, waschen mit *A. dest.*) unspezifisch angefärbt. Die gefärbten Größenstandards wurden mit Bleistift

markiert und die Membran entweder bei RT zwischen Filterpapier gelagert oder mit radioaktiv markierten Sonden hybridisiert (siehe 2.2.18)

1x TBE: 89 mM Tris, 89 mM Borsäure, 20 mM EDTA, pH 8,0

*crosslinking*-Lösung: 160 mM EDC, 130 mM 1-Methylimidazol, pH 8,0 mit HCl

Methylenblaulösung: 0,03 % (w/v) Methylenblau, 300 mM NaAc (pH 5,2)

### **2.2.21 Aufreinigung von Erythrozyten aus Mäuseblut**

Um Erythrozyten aus mit *P. berghei* infiziertem Mäuseblut zu isolieren wurden das Blut durch Cellulose-Säulen gefiltert. Die Säulen wurden selbst hergestellt, indem eine dünne Schicht Baumwollwatte am Boden einer Spritze festgestopft wurde. Auf die Watte wurde eine Schicht von ca. 3 cm Zellulosefasern (Sigma) gegeben. Die Zellulose wurde mit ca. 1 cm Glas-Granulat (212-300 µm, Sigma) überschichtet. Unmittelbar vor der Blutreinigung wurde die Säule mit 1,5 Säulenbettvolumen 1x PBS äquilibriert. Anschließend wurde das Blut auf die Säule gegeben. Nachdem das Blut in die Zellose-schicht eingesickert war, wurde die Säule mit 1x PBS überschichtet. Der Durchfluss aus der Säule wurde in einem 15 ml Zentrifugationsgefäß aufgefangen. Die Säule wurde solange mit 1x PBS überschichtet, bis das Blut die Säule passiert hatte bzw. das Gefäß gefüllt war. Die gereinigten Erythrozyten wurden anschließend zentrifugiert (1500 rpm, ohne Bremse) und die Parasiten wie in 2.2.32 beschrieben isoliert.

1x PBS: 14 mM NaCl, 1 mM Na<sub>2</sub>HPO<sub>4</sub>, 0,27 mM KCl, 0,18 mM KH<sub>2</sub>PO<sub>4</sub>, pH mit HCl auf 7,4

### **2.2.22 Isolation genomischer DNA aus Blutstadien von *P. berghei***

Für die Genotypisierung von transgenen *P. berghei* ANKA Parasiten durch PCR wurde die genomische DNA aus Blutstadien isoliert. Zur Isolation der Parasiten aus Mäuseblut wurde das Blut anfangs aufgereinigt (siehe 2.2.21) und die Erythrozyten anschließend durch Saponin (0,1 % (w/v) Saponin in 1x PBS) lysiert. Die freien Parasiten wurden durch Zentrifugation (1.500 x g, Heraeus Multifuge 1S-R, Thermo Fisher) sedimentiert. Um Rückstände des Saponins zu entfernen, wurde

der Zellschlag in 1 ml 1x PBS resuspendiert und erneut zentrifugiert (Tischzentrifuge 20 sec, 21.000 x g). Anschließend wurde die genomische DNA mit QIAamp DNA Blood Mini Kit (Qiagen) nach Herstellerangaben isoliert.

1x PBS: 14 mM NaCl, 1 mM Na<sub>2</sub>HPO<sub>4</sub>, 0,27 mM KCl, 0,18 mM KH<sub>2</sub>PO<sub>4</sub>, pH mit HCl auf 7,4

### 2.2.23 Anreicherung von Organellen aus *P. falciparum*

Die Anreicherung der Organellen erfolgte über differentielle Zentrifugation von *P. falciparum* NF54 Zelllysaten. Die Dauerkulturen (insgesamt 40 ml Kulturvolumen) der Blutstadien von *P. falciparum* wurden durch Zentrifugation sedimentiert und in 10 ml kaltem AIM-Puffer mit 0,075 % (w/v) Saponin resuspendiert um die Lyse der Erythrozyten auszulösen. Alle folgenden Schritte wurden bei 4°C bzw. im Eisbad durchgeführt. Die freigesetzten Parasiten wurden durch Zentrifugation sedimentiert (2800 x g, 13 Min.). Um Saponinreste zu entfernen, wurden die freien Parasiten zweimal in 10 ml kaltem AIM-Puffer resuspendiert und wie oben zentrifugiert. Nach dem zweiten Waschschrift wurden die Zellen in 10 ml kaltem MSE-Puffer resuspendiert und für die Stickstoff-Dekompression in die vorgekühlte *Cell disruption bomb* (Parr Instruments) transferiert. Die Zellen wurden für 20 Min. unter 80 bar Stickstoffdruck inkubiert. Die Probe wurde dann aus dem Druckbehälter abgelassen um die Dekompression auszulösen. Um ungebrochene Zellen und Zelltrümmer zu entfernen, wurde das Lysat für 10 Min. bei 2400 x g zentrifugiert. Die im Überstand enthaltenen, freien Organellen wurden anschließend durch Zentrifugation mit 23.000 x g (Sorval RC6 Zentrifuge, SS34 Rotor) gesammelt und in 500 µl MSE-Puffer resuspendiert. Aus dieser Fraktion wurde wie in 2.2.12 beschrieben die RNA extrahiert. Der Anreicherungs-erfolg wurde wie in 2.2.38 beschreiben ermittelt.

AIM-Puffer: 120 mM KCl, 20 mM NaCl, 10 mM PIPES, 1 mM MgCl<sub>2</sub>, 5 mM Glucose, pH 6,7 mit NaOH

MSE-Puffer: 225 mM Mannitol, 75 mM Saccharose, 0,1 mM EDTA, 5 mM PMSF, 1 % (w/v) BSA, 3 mM Tris-HCl pH 7,4



#### **2.2.24 Erstellung von cDNA-Bibliotheken für Hochdurchsatzsequenzierung**

Für die Identifikation von sRNA-Akkumulationen in den Organellen von *P. falciparum* wurden aus gesamt RNA und angereicherter RNA (siehe 2.2.23) mit dem Kit NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (NEB) cDNA-Bibliotheken für die Hochdurchsatzsequenzierung erstellt. Die Erstellung erfolgte nach Herstellerangaben mit 1 µg der jeweiligen RNA als Startmaterial. Die Größenselektion wurde nach der Amplifikation der cDNA-Bibliothek durchgeführt. dafür wurde die cDNA auf PAGE-Gelen aufgetrennt und die entsprechenden Bereiche ausgeschnitten. Die Elution erfolgte nach Herstellerangaben. Die Sequenzierung erfolgte kommerziell bei SMB (Berlin) auf einer MySeq-Plattform (Illumina).

#### **2.2.25 Bioinformatische Analyse von Hochdurchsatzsequenzierungsdaten**

Die Aufbereitung der in der Hochdurchsatzsequenzierung generierten Datensätze erfolgte durch mehrere bioinformatische Schritte. Anfangs wurden die Sequenzieradapter durch die Software *cutadapt* (Martin 2011) entfernt. Die beschnittenen Sequenzen wurden durch das Programm Bowtie 1.1.1 (Langmead et al. 2009) auf das Kern- und die Organellengenome von *P. falciparum* kartiert (Parameter: -a -best -strata -v 2). Durch diese Parameter werden nur die besten *Alignments* verwendet und ein Maximum von 2 *Missmatches* wird erlaubt. Gleichzeitig wurden nur einzigartige Sequenzen gewertet, Sequenzen, die auf mehrere Genome oder Position kartiert wurden, werden nicht berücksichtigt. Die verwendeten Genome stammen aus der NCBI Datenbank (Apicoplastengenom: DQ642846.1, Mitochondriengenom: DQ642845.1, Kerngenom: ASM276v1). Die erstellten Kartierungen wurden durch SAMtools (Li et al. 2009) bearbeitet und die Genomabdeckungen wurden durch BEDTools (Quinlan und Hall 2010) extrahiert. Die Visualisierung der Genomabdeckungen erfolgte über den Integrated Genome Browser (Freese et al. 2016). Zur Vervollständigung der mitochondrialen Genomannotation wurden die Sequenzinformationen der PlasmoDB und aus Feagin et al. 2012 verwendet. Für die Identifikation der sRNA-Akkumulationen wurden nur Fragmente genutzt, die mindestens 50 Einzelsequenzen am 5'-Ende zeigen.

### 2.2.26 Quantifizierung von Proteinlösungen

Die Konzentrationsbestimmung von Proteinen in Lösungen erfolgte durch den Protein-Assay (Bio-Rad). Parallel zu den Lösungen unbekannter Konzentration wurde eine BSA-Eichgerade erstellt. Die photometrische Absorptionsmessung erfolgte bei einer Wellenlänge von 595 nm im Novaspec II Spektrophotometer (Pharmacia).

### 2.2.27 Polyacrylamid-Gelelektrophorese (SDS-PAGE)

Die Analyse von Proteinproben erfolgte durch denaturierende Polyacrylamid-Gelelektrophorese. Die Proteine werden dabei durch Natriumdodecylsulfat (SDS, *sodium dodecyl sulfate*) denaturiert (SDS-PAGE, Laemmli 1970, Sambrook und Russell 2001). Neben der Qualitäts- und Mengenanalyse von Proteingemischen wurde die SDS-PAGE auch als Vorbereitung für den Transfer der Proteine auf eine Nitrocellulose-Membran (GE-Healthcare) genutzt (siehe 2.2.28). Die Proteinlösungen wurden mit 0,5 Vol. 3x Proteinladepuffer vermischt und für 5 min bei 95°C denaturiert. Die denaturierten Proben wurden auf ein 13 % SDS-PAGE-Gel aufgetragen. Der Gellauf erfolgte in vertikalen Gelkammern (Mini-Protean®, Bio-Rad) mit einer Spannung von 110 Volt. Parallel wurde ein Größenstandard (PageRuler™ Prestained Protein Ladder, Thermo Fisher) auf demselben Gel aufgetrennt. Als Laufpuffer in der Elektrophorese diente 1x SDS-Laufpuffer. Nach dem Gellauf wurden die Proteine entweder auf eine Membran transferiert (siehe 2.2.28) oder durch Coomassie-Färbung angefärbt und dokumentiert (ChemiDoc™ XRS+, Bio-Rad). Für die unspezifische Färbung durch Coomassie wurde das Gel zweimal in A. dest im Mikrowellenofen (40 s, maximale Stufe) erhitzt. Zur Färbung wurde das Gel in Coomassielösung nochmals erhitzt (20 s, maximale Stufe). Das Gel wurde dann bei RT inkubiert bis Banden im Gel sichtbar wurden. Zur Reduktion der Hintergrundfärbung wurde das Gel in A. dest bis zur gewünschten Intensität entfärbt.

Trenngellösung: 0,375 M Tris-HCl (pH 8,8), 13 % Polyacrylamid, 0,1 % (w/v) SDS, 0,1 % (w/v) APS, 0,05 % (v/v) TEMED

Sammelgellösung:	130 mM Tris-HCl (pH 6,8), 5 % Polyacrylamid, 0,1 % (w/v) SDS, 0,1 % (w/v) APS, 0,1 % (v/v) TEMED
1x SDS-Laufpuffer:	25 mM Tris-HCl, 192 mM Glycin, 0,1 % (w/v) SDS (pH 8,8)
3x Proteinladepuffer:	200 mM Tris-HCl (pH 6,8), 7 % (w/v) SDS, 30 % (v/v) Glycerol, 7,5 % (v/v) $\beta$ -Mercaptoethanol, 0,25 % (w/v) Bromphenolblau
Coomassielösung:	0,7 % (w/v) Coomassie® Brilliant Blue-G250, 35mM HCl

### 2.2.28 Proteintransfer auf Nitrocellulose-Membranen (*Western-blot*)

Für die Detektion von Proteinen wurden diese auf eine Nitrocellulose-Membran transferiert und mit einem spezifischen Antikörper detektiert. Die Proteinlösungen wurden hierzu mit einer SDS-PAGE aufgetrennt und anschließend elektrophoretisch auf eine Nitrocellulose Membran (GE-Healthcare) transferiert (Sambrook und Russell 2001). Der Transfer wurde im Tank-*blot* (Mini Trans-Blot Cell, Bio-Rad) in 1x Transferpuffer durchgeführt (Spannung: 100 V). Alternativ wurde der Transfer über Nacht bei 4°C und 36V durchgeführt. Durch die Färbung der Membran mit Ponceau-Lösung (10 min bei RT) und mehrfachem Waschen mit *A. dest.* wurde der Transfererfolg überprüft und dokumentiert (Klein et al. 1995). Die Membran wurde entweder bei RT zwischen Filterpapier gelagert oder direkt in der Immundetektion verwendet.

Die immunologische Detektion von Proteinen auf Nitrocellulose-Membranen erfolgte durch spezifische Antikörper. Nach dem Proteintransfer wurden zuerst freie Stellen der Membran durch Blockierlösung (30 - 60 min, leichtes Schwenken) abgedeckt. Anschließend wurde der entsprechende Antikörper in 15 ml Antikörperlösung verdünnt (siehe 2.1.7) und 1 Stunde unter leichtem Schwenken mit der Membran inkubiert. Nach 3 achtminütigen Waschschritten mit 1x TBST wurde die Membran mit dem passenden sekundären Antikörper inkubiert. Dieser wurde in 1x TBST verdünnt (Verdünnung siehe 2.1.7). Nach 45 min Inkubation erfolgten wieder drei achtminütige Waschschrritte (1x TBST). Die direkte Detektion erfolgte durch an den sekundären Antikörper gekoppelte HRP (*horseradish*

*peroxidase*). Zur Aktivierung der HRP-Reaktion wurde die Membran für 1 – 5 min in Detektionslösung geschwenkt. Die Detektion erfolgte im ChemiDoc XRS+ (Bio-Rad).

10x Transferpuffer:	25 mM Tris, 192 mM Glycin, 20 % (v/v) Methanol
1x Ponceau-Lösung:	0,1 % (w/v) Ponceau S, 1 % (v/v) Essigsäure
1x TBST:	50 mM Tris-HCl (pH 7,5), 150 mM NaCl, 0,1% (v/v) Tween20
Blockierlösung:	4 % (w/v) Trockenmilchpulver in 1x TBST
Antikörperlösung:	2 % (w/v) Trockenmilchpulver in 1x TBST
Detektionslösung:	0,45 mM Coumarat, 1,25 mM Luminol, 0,01 % (w/v) H <sub>2</sub> O <sub>2</sub> in 100 mM Tris-HCl (pH 8,5)

### 2.2.29 Heterologe Expression von Proteinen

Für die heterologe Expression von rekombinanten Proteinen in *E. coli* wurde mit Änderungen das pMAL™ Protein Fusion and Purification System (NEB) genutzt. Die Fusion eines Proteins mit einem Maltose-Binde-Protein (MBP) konnte die Löslichkeit von in *E. coli* exprimierten Proteinen drastisch erhöhen (Kapust und Waugh 1999). Der Expressionsvektor pMAL-c5X wurde so modifiziert, dass zwischen MBP und dem rekombinanten Protein eine *Tabacco Etch Virus* (TEV) Erkennungssequenz eingesetzt wurde. Diese wird spezifisch von der TEV-Protease erkannt (Expressionsvektor pMAL-TEV, Modifikation von Prof. Dr. A. Barkan, University of Oregon, USA). Das verwendete Plasmid enthält außerdem bereits ein Fusionsprotein aus MBP:CDC42 (*cell division control protein 42*). CDC42 stammt aus *Drosophila melanogaster* (NCBI: NP\_001245762.1) und wurde als Kontrolle in *in vitro* RNA-Bindestudien verwendet.

#### 2.2.29.1 Erstellung der Expressionskonstrukte

Die kodierenden Sequenzen der *P. berghei* Proteine wurden mittels PCR amplifiziert (siehe 2.2.4) und aufgereinigt (siehe 2.2.9). Durch die verwendeten

Oligonukleotide wurden zusätzlich Erkennungssequenzen für die Restriktionsenzyme *Bam*HI (5'Ende, NEB) und *Sall* (3'Ende, NEB) in das PCR-Produkt eingefügt. Der Verdau der PCR-Produkte und der Plasmid-DNA mit den Restriktionsenzymen erfolgte nach Herstellerempfehlung. Bei der Plasmid DNA wurde zusätzlich während der Reaktion eine alkalische Phosphatase (FastAP, Thermo Fisher) zugesetzt. Diese spaltet freie Phosphatgruppen von den 5'-Enden ab. Nach einer weiteren Aufreinigung nach 2.2.9 wurden die geschnittenen Fragmente mit dem verdauten Plasmid in einer T4-Ligasereaktion eingesetzt (siehe 2.2.10). Mit den ligierten DNA-Molekülen wurden wie in 2.2.5 beschrieben *E. coli* TOP10 Zellen transformiert und auf LB-Agarplatten mit Carbenicillin (100 µg/ml) ausplattiert. Die jeweiligen Klone wurden mittels Kolonie-PCR analysiert (2.2.6). Klone, die eine Insertion der erwarteten Größe zeigten, wurden nochmals über Nacht kultiviert und die Plasmide isoliert (2.2.2.2; 2.2.9). Die korrekte Sequenz wurde mittels Sanger-Sequenzierung verifiziert. Mit den fertigen Expressionskonstrukten wurden *E. coli* RIPL Zellen transformiert (2.2.5) und für die Proteinexpression genutzt.

#### **2.2.29.2 Expression der rekombinanten Proteine in *E. coli***

Die Expression der rekombinanten Fusionsproteine erfolgte wie vom Hersteller beschrieben. Eine einzelne Kolonie der transformierten *E. coli* RIPL Zellen wurde über Nacht in LB-Flüssigkultur (2.2.2.2) mit 100 µg/ml Carbenicillin angezogen. 500 ml LB-Flüssigkultur wurden am Folgetag mit 5 ml der Übernachtskultur inokuliert und bei 37°C inkubiert. Das Wachstum der Kultur wurde durch photometrische Absorptionsmessung bei 600 nm Wellenlänge überwacht (Novaspec II Spektrophotometer, Pharmacia). Bei einer optischen Dichte von 0,5 wurden dem Kulturmedium IPTG (Isopropyl-β-D-thiogalactopyranosid) zu einer Endkonzentration von 0,3 mM zugefügt und die Zellen über Nacht bei 37°C kultiviert. Am Folgetag wurden die Zellen durch Zentrifugation geerntet (4000 x g, 4°C, 20 min). Der Überstand wurde verworfen und die Zellen in 12 ml Säulen-Puffer mit 0,2 mg/ml Lysozym resuspendiert. Zum Schutz der Proteine gegen die *E. coli* eigenen Proteasen wurden zusätzlich Proteinaseinhibitoren verwendet (cOmplete™, *Protease Inhibitor Cocktail*, Roche). Der Zellaufschluss durch Lysozym wurde durch Ultraschallbehandlung unterstützt. Diese wurde bei 4°C im

Kälteraum durchgeführt, mit 6 x 30 s Puls. Um Zelltrümmer zu entfernen, wurden die Zelllysate anschließend für 20 min und 20.000 x g zentrifugiert. Die Proteinkonzentration der Überstände wurde nach Bradford bestimmt (2.2.26). Außerdem wurden Proben für die immunologische Proteinanalyse abgenommen (2.2.28).

### **2.2.29.3 Reinigung von rekombinanten Fusionsproteinen mittels Affinitätschromatographie**

Für die Affinitätschromatographie wurden 3 ml Amylose-Suspension (NEB) fünfmal mit Säulen-Puffer gewaschen. Der Puffer aus dem letzten Waschschrift wurde abgenommen und die Amylose mit den Überständen der Zelllysate gemischt. Die Bindung der Proteine an die Amylose erfolgte bei 4°C über Nacht unter leichter Rotation. Am Folgetag wurde die Suspension in eine Säulenform gegossen und mit 12 Säulenvolumen Säulen-Puffer gewaschen. Die Elution der Proteine erfolgte mit Säulen-Puffer mit 10 mM Maltose. Die Eluate wurden in 1 ml Fraktionen aufgefangen und eine Probe jeder Fraktion durch SDS-PAGE und *Western-blot* (2.2.27) mit Anti-MBP-Antikörpern analysiert. Die Fraktionen, in denen in der Immundetektion Fusionsproteine erkennbar waren, wurden vereinigt und durch Zentrifugation durch eine Membran (Amicon ultra, 0,5 ml 50KDa Ausschlussgröße) zu einem Endvolumen von 150 µl konzentriert. Die Proteine wurden anschließend mit 150 µl 50 % Glycerol gemischt und bei -20°C gelagert.

Säulen-Puffer:                    20 mM Tris-HCl (pH 7,4), 200 mM NaCl, 1mM EDTA, 1 mM DTT

### **2.2.30 RNA-Bindungsstudie von rekombinanten Proteinen**

Um das RNA-Bindepotential von rekombinanten Fusionsproteinen zu untersuchen, wurden *in vitro* Bindestudien durchgeführt. Hierzu wurden zum einen die Fusionsproteine rekombinant in *E. coli* exprimiert (2.2.29) und zum anderen das mitochondriale Transkriptom durch *in vitro* Transkription hergestellt (2.2.15). Von den Fusionsproteinen wurden 580 pmol eingesetzt, von jedem RNA-Fragment 6 µg. Die Proteine wurden für 20 min bei RT in 150 µl Säulen-Puffer inkubiert. Die strangspezifischen RNA-Fragmente wurden in 50 µl Säulen-Puffer mit 2 mM EDTA

gegeben und für 3 min bei 95°C denaturiert. Anschließend wurden die Fragmente auf Eis gekühlt. Von den denaturierten Fragmenten wurden Proben für die Gelanalyse genommen (jeweils 400 ng) und die RNA-Lösung den Proteinen hinzugefügt. Die rekombinanten Proteine wurden zusammen mit den RNAs für 20 min bei RT unter Rotation inkubiert. Parallel wurden 200 µl Amylose-Suspension fünfmal mit Säulen-Puffer gewaschen. Nach dem letzten Waschschrift wurde die Amylose in der Protein-RNA-Lösung resuspendiert und für weitere 20 min inkubiert. Die Amylose wurde zweimal mit 300 µl Säulen-Puffer gewaschen. Die Elution der Proteine erfolgte durch inkubation der Amylose in 200 µl Säulen-Puffer mit 10 mM Maltose für 10 min. Von jedem Schritt des Experiments wurden 10 % für die Immundetektion abgenommen. Aus der Elutionsfraktion wurde wie in 2.2.12 beschrieben die RNA extrahiert und auf denaturierenden RNA-Agarosegelen aufgetrennt (2.2.16).

Säulen-Puffer:                    20 mM Tris-HCl (pH 7,4), 200 mM NaCl, 1mM EDTA, 1 mM DTT

### **2.2.31 Giemsa Färbung für die Bestimmung der Parasitämie**

Um die Anzahl der infizierten Erythrozyten in Blutproben zu bestimmen, wurde eine Giemsa-Färbung durchgeführt. Dafür wurde eine kleine Blutmenge aus dem Schwanz einer mit *P. berghei* infizierten Maus abgenommen und auf einem Objektträger ausgestrichen. Nach einer kurzen Fixierung durch Methanol wurde der Objektträger für 10-15 min in 10 % Giemsa-Lösung (VWR) inkubiert. Überschüssige Färbelösung wurde mit *A. dest.* abgewaschen. Die mit Parasiten infizierten Blutzellen und die nicht infizierten Blutzellen wurden im Lichtmikroskop gezählt. Die Parasitämie wurde durch den Quotienten aus der Anzahl der Parasiten und der die Gesamtzahl der gezählten Erythrozyten bestimmt.

### **2.2.32 Isolation von Parasiten aus infiziertem Blut**

Für die Isolation von Parasiten aus Blutkulturen bzw. Mäuseblut wurden die Blutzellen durch Zentrifugation gesammelt (450 x g ohne Bremse, Heraeus Multifuge 1S-R, Thermo Fisher). Der Überstand wurde verworfen und die Blutzellen in 1x PBS mit 0,1 % (w/v) Saponin resuspendiert. Die freien Parasiten wurden

ebenfalls durch Zentrifugation gesammelt (1500 x g mit Bremse) und der Überstand verworfen. Zum Entfernen von verbliebenem Saponin wurde der Zellschlag in 1 ml 1x PBS resuspendiert und in ein Mikroreaktionsgefäß überführt. Die Parasiten wurden in einer Tischzentrifuge (Heraeus Fresco Mikrozentrifuge, Thermo Fisher) für 10 s bei 21.000 x g erneut zentrifugiert. Die freien Parasiten wurden dann in kleinem Volumen 1x PBS resuspendiert.

1x PBS: 14 mM NaCl, 1 mM Na<sub>2</sub>HPO<sub>4</sub>, 0,27 mM KCl, 0,18 mM KH<sub>2</sub>PO<sub>4</sub>, pH mit HCl auf 7,4

### **2.2.33 Transfektion von *Plasmodium berghei* ANKA**

#### **2.2.33.1 Anreicherung von Schizonten für die Transfektion**

Die stabile Integration von Fremd-DNA in das Genom von *P. berghei* ANKA Parasiten wurde basierend auf Janse et al. 2006 mit Vektoren nach Kooij et al. 2012 durchgeführt. Zur Vorbereitung der Transfektion wurden NMRI-Mäuse mit Wildtyp Parasiten von *P. berghei* ANKA infiziert. Bei einer Parasitämie von etwa 2 % wurde den mit Isofluran betäubten Mäusen das Blut abgenommen. Die für die Herzpunktion verwendeten Kanülen wurden vorher mit Heparin behandelt. Das Blut wurde anschließend in 10 ml vorgewärmtem Transfektionsmedium (37°C) mit 200 U/ml Heparin (Ratiopharm) resuspendiert. Nach einer Zentrifugation (8 min, 1000 rpm, ohne Bremse) wurde der Überstand verworfen und die Zellen in 25 ml vorgewärmtem Transfektionsmedium resuspendiert. Die Lösung wurde in einem Erlenmeyerkolben in weiteren 125 ml vorgewärmtem Transfektionsmedium (37°C) transferiert und in einem Gasinkubator (Mytron) für 16-18 Stunden bei leichtem Schwenken (77 rpm) inkubiert. Die Atmosphäre des Inkubators bestand dabei aus 10 % Sauerstoff, 5 % Kohlenstoffdioxid und 85 % Stickstoff. Die Parasiten entwickeln sich *in vitro* zu Schizonten, die Freisetzung der Merozoiten aus den Erythrozyten ist jedoch inhibiert.

Transfektionsmedium: 160 ml RPMI 1640 (PAN-Biotech), 40 ml Kälberserum (USA, Inaktiviert 30 min bei 56°C), 60 µl Gentamycin (Life Technologies)



### 2.2.33.2 Isolation der Schizonten

Eine Probe der Übernachtskultur wurde auf einem Objektträger ausgestrichen und mit Giemsa-Lösung (VWR) gefärbt um die Anreicherung der Schizonten zu bewerten (siehe 2.2.31). Um die Schizonten zu isolieren, wurde die Kultur auf vier 50 ml Zentrifugationsgefäße aufgeteilt (je 35 ml). Die Kulturen wurden vorsichtig mit 10 ml Nycodenz-Lösung unterschichtet und für 30 min bei 1500 rpm zentrifugiert (ohne Bremse). Anschließend wurde die bräunliche Schicht über der Nycodenz-Schicht mit einer Pasteur-Pipette abgenommen und in zwei frische 50 ml Gefäße transferiert. Die Schizonten wurden durch Zentrifugation (1500 rpm, 10 min) gesammelt. Für jeden geplanten Transfektionsansatz wurde den Zellen 1 ml Medium zugesetzt. Die resuspendierten Zellen wurden auf 1,5 ml Zentrifugationsgefäße aufgeteilt (jeweils 1 ml) und für 5 s bei voller Geschwindigkeit in der Tischzentrifuge sedimentiert.

Nycodenz-Lösung: 55 % (v/v) Nycodenz in 1 x PBS

### 2.2.33.3 Transfektion von *P. berghei* mittels Elektroporation

Die Integration der erstellten Konstrukte für die Deletion bzw. Markierung von Proteinen erfolgte durch doppelte homologe Rekombination der Konstrukte mit dem Genom von *P. berghei* ANKA. Die Konstrukte wurden vor der Transfektion durch den Verdau mit passenden Restriktionsendonukleasen (für alle Konstrukte *Sall* (NEB)) linearisiert und wie in 2.2.14 beschrieben präzipitiert. Die Konstrukte wurden anschließend in 10 µl *A. bidest.* resuspendiert. Für die Transfektion wurde das Human T Cell Nucleofector™ Kit (Lonza) verwendet. Die linearisierten Konstrukte wurden kurz vor der Transfektion mit 100 µl der *human T-cell nucleofactor*-Lösung des Transfektionskits versetzt. In dieser Lösung wurden die gesammelten Schizonten resuspendiert und in eine Elektroporationsküvette überführt. Die Elektroporation erfolgte im Elektroporator (Nucleofector 2b Device, Amaxa, Programm U33). Anschließend wurden dem Ansatz unverzüglich 50 µl Transfektionsmedium hinzugefügt und intravenös in eine NMRI-Maus injiziert. Die Selektion mit Pyrimethamin erfolgte am Folgetag durch das Trinkwasser der Mäuse.

### 2.2.34 Fluoreszenzmarkierung von Proteinen in *P.berghei*

Die genetische Markierung von Proteinen im Genom von *P. berghei* erfolgte über doppelte homologe Rekombination. Das verwendete Plasmid (pBAT1-G) basierte auf dem pBAT-SIL6 Vektorsystem (Kooij et al. 2012). Eine schematische Darstellung des Vektors ist im Ergebnisteil (siehe Abbildung 16) gezeigt. Durch diesen Vektor werden die Proteine mit einem fluoreszenten mCherry sowie einem biochemisch nutzbaren, dreifachen C-Myc Proteintag markiert. Neben den Komponenten der Markierung enthält das Plasmid einen bakteriellen Selektionsmarker (Resistenz gegen Carbenicillin) und einen Selektionsmarker im Maussystem (DHFR, Resistent gegen Pyrimethamin). Die im Plasmid enthaltene GFP-Expressionskasette wurde in Vorarbeiten in der Arbeitsgruppe Matuschewski gegen ein hoch-exprimiertes GFP mit mitochondrialer Lokalisation ausgetauscht (Matz et al. 2018). Dies ermöglicht die Identifikation des Mitochondriums in mikroskopischen Anwendungen.

Für die genetische Markierung wurden die genomischen Sequenzen stromauf des STOP-Codons sowie der 3'-UTR Bereich des zu markierenden Proteins durch PCR amplifiziert. Durch die verwendeten Oligonukleotide wurden gleichzeitig die Erkennungssequenzen für Endonukleasen an die Enden der Amplifikate angefügt. Die 5'-homologe Sequenz wurde mit *EagI* und *HpaI* kloniert, der 3'-Bereich mit *XhoI* und *HindIII* (alle NEB). Die 5'-homologe Sequenz lag dabei im gleichen Leseraster wie die Sequenz des Fluorophors. Der Vektor und die amplifizierten und gereinigten DNA-Fragmente wurden mit den passenden Restriktionsendonukleasen verdaut und anschließend wie in 2.2.9 beschrieben gereinigt. Die Ligation erfolgte über Nacht wie in 2.2.10 beschrieben. Mit 2-5 µl der Ligrationsreaktion wurden *E. coli* TOP10 Zellen transformiert (siehe 2.2.5). Plasmide mit erfolgreich ligierten DNA-Fragmenten wurden durch Sanger-Sequenzierung validiert. Für die Transfektion wurden 15 µg des Plasmids durch den Verdau mit *SalI* (NEB) linearisiert und mit Ethanol präzipitiert (siehe 2.2.14). Die Transfektion von *P. berghei* ANKA erfolgte wie in 2.2.33 beschrieben.

Die transgenen Parasiten wurden für 5-10 Tage in NMRI-Mäusen kultiviert. Die mikroskopische Analyse erfolgte wie in 2.2.36 beschrieben. Für die Analyse des Markierungserfolges wurde die genomische DNA der Parasiten wie in 2.2.22

beschrieben isoliert und durch PCR analysiert. Die Position der für die Genotypisierung verwendeten Oligonukleotide ist im Ergebnisteil exemplarisch gezeigt (siehe Abbildung 16).

### **2.2.35 Deletion von Proteinen in *P. berghei***

Die Deletion von ausgewählten HPR-Proteinen erfolgte durch doppelte homologe Rekombination. Das verwendete Deletionskonstrukt (pBAT-M6) basiert auf dem pBAT-SIL6 Vektorsystem (Kooij et al. 2012). Auf dem Plasmid ist ein cytosolisch-lokalisierter Fluoreszenzmarker (mCherry) enthalten, der die Identifikation und Isolation von transgenen Parasiten durch FACS (Fluorescence activated cell sorting) ermöglicht. Für die chemische Selektion der Parasiten in den Blutstadien ist ein Selektionsmarker (DHFR) enthalten, der die Resistenz gegen Pyrimethamin vermittelt enthalten.

Für die Deletion ausgewählter Proteine wurden die Sequenzen stromauf und stromab des entsprechenden Genlokus mittels PCR amplifiziert. Durch die verwendeten Oligonukleotide wurden gleichzeitig Erkennungssequenzen für Restriktionsendonukleasen eingefügt. Zum Verdau der 5'-homologen Sequenz wurden *EagI* und *HpaI*, für die 3'-homologe Sequenz *XhoI* und *HindIII* bzw. *Clal* verwendet. Der Vektor wurden parallel mit den entsprechenden Nukleasen verdaut. Die verdaute DNA wurde anschließend wie in 2.2.9 beschrieben aufgereinigt. Nach der Aufreinigung wurden die verdauten DNA-Fragmente und die Plasmid-DNA mit T4-DNA-Ligase (NEB) über Nacht ligiert (siehe 2.2.10). Die Ligase wurde durch Hitze inaktiviert (20 min, 65°C). Die Reaktion wurde für die Transformation von *E. coli* genutzt (siehe 2.2.5). Gewachsene Klone wurden mittels Kolonie-PCR (2.2.6) analysiert. Klone mit Insertionen der gesuchten Größe wurden erneut angezogen (2.2.2.2). Die Sequenzintegrität wurde mittels Sanger-Sequenzierung bestätigt.

Für die Transfektion von *P. berghei* ANKA Parasiten wurden 15 µg des Deletionsplasmids mit *Sall* (NEB) linearisiert. Durch den Verdau mit diesem Enzym wird das bakterielle Vektorrückrad und der für die Deletion im Parasiten wichtige Teil getrennt. Nach dem Verdau wurde die DNA präzipitiert (siehe 2.2.14) und in 10 µl *A. bidest.* aufgenommen. Die Transfektion erfolgte wie in 2.2.33 beschrieben. Für die Analyse des Deletionserfolges mittels PCR wurde die genomische DNA der

Parasiten wie in 2.2.22 beschrieben isoliert. Die Position der für die Genotypisierung verwendeten Oligonukleotide ist im Ergebnisteil exemplarisch gezeigt (siehe Abbildung 18).

### **2.2.36 Fluoreszenzmikroskopie**

Die mikroskopische Analyse von fluoreszent markierten Proteinen erfolgte in lebenden Zellen mit einem Zeiss AxioObserver Z1 Epifluoreszenzmikroskop. Hierfür wurde ein Blutstropfen von mit transgenen *P. berghei* infizierten Mäusen auf einem Objektträger gegeben und mit 0,1 µl Hoechst 33342 (Thermo Fisher) gemischt. Die unspezifische DNA Färbung durch Hoechst 33342 ermöglicht die mikroskopische Identifikation von Zellkernen. Die Zellen wurden nicht fixiert. Die Bildbearbeitung erfolgte mit ImageJ Fiji.

### **2.2.37 Reverse Transkription**

Um cDNA-Bibliotheken für die quantitative PCR zu erstellen, wurde 1 µg RNA mittels SuperScript™ III Reverse-Transkriptase (Invitrogen) in DNA umgeschrieben. Die Nutzung erfolgte nach Herstellerangaben. Als Primer dienten Hexa- und Nona-Oligonukleotide (6/9) mit zufälliger Sequenz. Nach erfolgter Transkription wurde die cDNA als Matrize in quantitativen PCR-Experimenten verwendet.

### **2.2.38 Quantitative PCR**

Die Bestimmung des Anreicherungs Erfolges der Organellen-Präparationen aus *P. falciparum* erfolgte über quantitative RT-PCR (qRT-PCR). Für die Erstellung der cDNA-Bibliotheken wurde RNA aus den verschiedenen Fraktionen der Präparation isoliert (siehe 2.2.23 und Abbildung 6). Die isolierten RNA wurde wie in 2.2.37 beschrieben in cDNA umgeschrieben. Die qPCR wurde im 7500 Fast Real-Time PCR System (Applied Biosystems) durchgeführt. Die Planung der Experimente erfolgte mit der entsprechenden Software (7500 Fast Software) durchgeführt. Es wurden jeweils 3 technische Replikate für jede cDNA erstellt, jedes mit einem Endvolumen von 15 µl. Gleichzeitig wurde für jedes Oligonukleotidpaar eine Kontrolle ohne cDNA mitgeführt. Die Reaktion wurde nach Herstellerangaben zusammengestellt, wobei der SYBR™ Green PCR Master Mix (Applied Biosystems) verwendet wurde. Von den cDNA-Bibliotheken wurden jeweils 1,5 µl in die Reaktion

eingesetzt. Die Auswertung der Ergebnisse erfolgte mit der oben genannten Software.

### 3. Ergebnisse

#### 3.1 Vorbemerkungen

Die im Folgenden dargestellten Ergebnisse sind teilweise im Journal *Nucleic Acid Research* publiziert. Die Publikation hat den Titel:

**Identification of clustered organellar short (cos) RNAs and of a conserved family of organellar RNA-binding proteins, the heptatricopeptide repeat proteins, in the malaria parasite.** Arne Hillebrand, Joachim M. Matz, Martin Almendinger, Katja Müller, Kai Matuschewski, Christian Schmitz-Linneweber, *Nucleic Acids Research*, Volume 46, Issue 19, 2. November 2018, Pages 10417–10431

Abbildungen die entweder aus der Publikation oder darauf basierend verändert in dieser Arbeit verwendet wurden, sind durch die Zitation „Hillebrand et al 2018“ gekennzeichnet.

Einige Teile dieser Arbeit wurden in Zusammenarbeit mit Studierenden und Mitarbeitern unter meiner Betreuung durchgeführt. MSc Martin Almendinger war an der Lokalisation zweier HPR-Proteine beteiligt. Sabrina Tetzlav war an der Erstellung der Deletionsmutanten beteiligt. Die bioinformatische Analyse der Hochdurchsatzsequenzierung wurde mit der Unterstützung von Gongwei Wang durchgeführt. Die Modellierung der HPR-Proteine wurde von Prof. Dr. Christian Schmitz-Linneweber durchgeführt. Sämtliche Mausearbeiten wurden von Dr. Joachim M. Matz und Dr. Katja Müller durchgeführt. Alle weiteren Experimente wurden von mir selbst geplant und durchgeführt.

#### 3.2 Sequenzierung von organellären kurzen RNAs aus *P. falciparum*

In höheren Pflanzen wurde gezeigt, dass Proteine mit repetitiven helikalen Elementen wie PPR-Proteine ihre Bindestelle an der RNA vor Exonukleasen beschützen und dabei eine charakteristische Akkumulation einer kurzen RNA verursachen (z B. Ruwe et al. 2016). Um zu untersuchen, ob es eine ähnliche

protektive Aktivität von helikalen Proteinen in apicomplexen Parasiten gibt, wurden die zellulären kurzen RNAs aus angereicherten Organellen sequenziert. Hierzu musste zunächst ein Protokoll etabliert werden, um Mitochondrien und Apicoplasten anzureichern. Die notwendigen Parasiten von *P. falciparum* NF54 wurden von der Arbeitsgruppe „Vektorbiologie“ von Dr. Elena A. Levashina (MPIIB, Berlin) kultiviert und zur Verfügung gestellt.

### **3.2.1 Die Organellen von *P. falciparum* können durch differentielle Zentrifugation angereichert werden**

Die Anreicherung spezifischer Zellbestandteilen aus Organismen wird in der Regel durch die Zentrifugation der Zellysate in Dichtegradienten erreicht. Für Organellen von *P. falciparum* wurde bereits ein Anreicherungsprotokoll etabliert (Kobayashi et al. 2007). Für den Zellaufschluss wird in diesem Protokoll eine „cell disruption bomb“ (Parr Instruments Company) verwendet. Die Zellen werden dabei durch Stickstoff-Dekompression aufgeschlossen, wodurch Zellbestandteile wie Organellen größtenteils intakt freigesetzt werden (Takashima et al. 2001). Basierend auf diesen Studien wurde in dieser Arbeit eine Anreicherungsmethode für die Organellen von *P. falciparum* Blutstadien etabliert. Zur Bestimmung des Anreicherungs Erfolges wurde durch reverse Transkription und quantitative-PCR (qRT-PCR) die Menge der plastidären mRNA *sufB* (PlasmoDB: PF3D7\_API04700) in Relation zur 18S rRNA (PlasmoDB: PF3D7\_1148600) des Kerns bestimmt. Der schematische Ablauf der Anreicherung ist in Abbildung 6 A gezeigt (siehe auch 2.2.23). Nach der RNA-Extraktion aus den Proben wurden durch reverse Transkription cDNA Bibliotheken generiert, in denen durch quantitative PCR das relative Verhältnis der beiden Transkripte bestimmt wurde. Als Ausgangswert wurde dabei das aufgeschlossene Material vor der ersten Zentrifugation genutzt (Fraktion Input, siehe Abbildung 6 ). Das Verhältnis der beiden RNAs in dieser Fraktion wurde zur Normalisierung als 1 angenommen und basierend darauf die Folgefraktionen verglichen. Die Auswertung der qPCR zeigte, dass mit dem entwickelten Protokoll eine etwa achtfache Anreicherung der RNAs des Apicoplasten erreicht werden konnte (siehe Abbildung 6 B, Fraktion Pellet 2). Die Fraktion „Pellet 2“ wird im Folgenden als angereicherte Fraktion bezeichnet. Sie wurde für die Erstellung von Sequenzierbibliotheken genutzt (siehe 3.2.2). Bereits im Pellet 1 ist eine vierfache

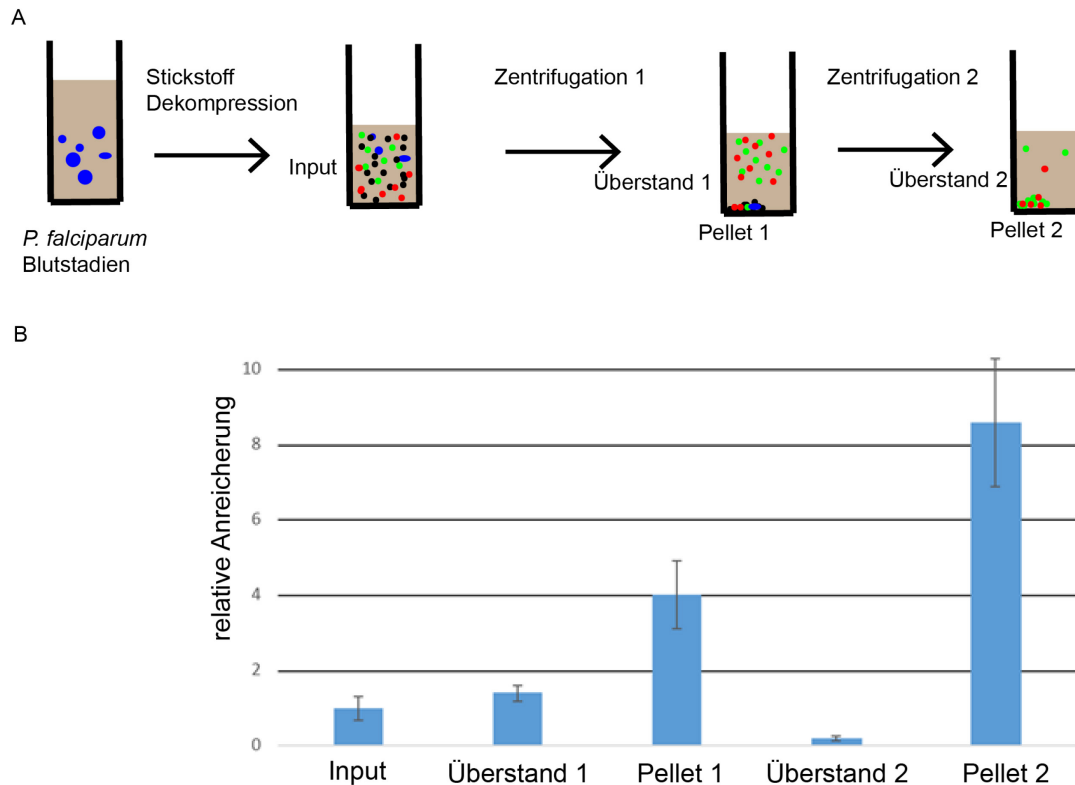


Abbildung 6: Anreicherung der Organellen von *P. falciparum* Blutstadien. **A)** Nach der Lyse der Erythrozyten durch Saponin wurden die freigesetzten Parasiten gewaschen und in die *cell disruption bomb* transferiert. Nach der Inkubation unter Stickstoffdruck erfolgte der Zellaufschluss durch Dekompression. Die freigesetzten Zellbestandteile (blaue Punkte = intakte Parasiten, schwarze Punkte = Zellkerne, rote Punkte = Mitochondrien, grüne Punkte = Apicoplasten) wurden anschließend durch Zentrifugationen getrennt. Durch die Zentrifugation 1 wurden Zelltrümmer, intakte Zellen und Zellkerne sedimentiert (Pellet 1). Aus dem Überstand 1 wurden anschließend in der Zentrifugation 2 die Mitochondrien und Apicoplasten sedimentiert (Pellet 2). **B)** Die Anreicherung des Apicoplasten wurde durch quantitative PCR bestimmt, bei der das Verhältnis der apicoplastidären *sufB* mRNA zur cytosolischen 18S rRNA ermittelt wurde. Der Input wurde zur Normalisierung der Ergebnisse genutzt. In beiden Zelniederschlägen ist eine Anreicherung der *sufB* mRNA zu erkennen. Die extrahierte RNA aus dem Pellet 2 wurde für die Sequenzier- und für *Northern-Blot* Experimente genutzt (Hillebrand et al. 2018).

Anreicherung zu erkennen, obwohl die Organellen hier im Idealfall im Überstand verbleiben sollten (siehe Abbildung 6, Fraktion Pellet 1). Diese kann möglicherweise durch nur partiell aufgeschlossene Parasiten erklärt werden, bei denen das Cytosol mit der 18S rRNA aus den Zellen austritt, ohne dass eine vollständige Zelllyse eintritt. Die Organellen werden bei der anschließenden Zentrifugation mit den



Zelltrümmern vom Überstand abgetrennt und angereichert. Durch eine Wiederholung des Zellaufschlusses könnte dieses Problem minimiert werden; da die festgestellte Anreicherung ausreichend für die folgenden Experimente war, wurde keine Optimierung des Zellaufschlusses durchgeführt.

### **3.2.2 Erstellung von Sequenzierbibliotheken und Hochdurchsatzsequenzierung kurzer RNAs aus *P. falciparum***

Für die Hochdurchsatzsequenzierung (HTS = *high throughput sequencing*) wurde RNA aus der angereicherten Fraktion (enRNA), sowie aus der unbehandelten zellulären RNA (totRNA) extrahiert und genutzt, um Sequenzierbibliotheken herzustellen. Die beiden RNAs wurden vor der Erstellung der Bibliotheken mittels Gelelektrophorese auf Qualität und Degradation überprüft. In beiden RNA-Präparationen konnten klare Banden für die cytosolischen rRNAs identifiziert werden. Dies lässt auf einen geringen Degradationsgrad schließen. Bei stärkerer RNA-Degradation erscheinen die rRNA-Banden in Agarosegelen unscharf. Die Erstellung der Sequenzierbibliotheken erfolgte mit dem Kit NEBNext® Small RNA Library Prep (NEB) bei der jeweils 1 µg RNA eingesetzt wurde. Die Größenselektion der Sequenzierbibliotheken erfolgte nach der Amplifikation der cDNA. Die Sequenzierbibliotheken wurden mittels Gelelektrophorese in einem PAGE-Gel aufgetrennt und mit Ethidium-Bromid visualisiert. Es wurden die Bereiche aus dem Gel ausgeschnitten die mit RNAs zwischen 17 und 100 nt korrespondieren und cDNA extrahiert. Die Sequenzierung wurde kommerziell durchgeführt, wobei die Sequenzierlänge auf maximal 100 bp beschränkt wurde.

Die Sequenzierung generierte zwei Datensätze. Die Sequenzierung der totRNA lieferte insgesamt 12.726.675 Einzelsequenzen (*reads*); die angereicherte enRNA Sequenzierung generierte 10.594.281 *reads*. Das Zuordnen der Sequenzen zu den Organellengenomen (verwendete Genome: Apicoplast DQ642846.1; Mitochondrium, DQ642845.1) bzw. zu den Chromosomen des Kerngenoms wurde, wie in 2.2.25 beschrieben durchgeführt. Aus der totRNA-Sequenzierung konnten 4.111 (0,03 % der gesamten Sequenzen) einzigartige *reads* dem Apicoplastengenom und 1.073.135 (8,43 %) dem Mitochondriengenom zugeordnet werden. In der enRNA wurden dem Plastidengenom 23.462 (0,22 %) und dem mitochondrialen Genomen 2.048.814 (19,34 %) *reads* zugeordnet. Für den

Apicoplasten ergibt sich daraus eine siebenfache Anreicherung und für das Mitochondrium eine etwa zweifache Anreicherung im Vergleich mit dem Ausgangsmaterial.

Besonders hervorzuheben ist, dass in der zellulären totRNA 9 % der Sequenzen mitochondriellen Ursprungs sind und nur ein verschwindend geringer Teil ihren Ursprung im Apicoplasten hat. In vergleichbaren Experimenten mit pflanzlicher RNA entfallen nur ca. 1,6 % der gesamten *reads* auf das mitochondriale Genom, während dem Genom der Plastiden 20,4 % zugeordnet werden können (siehe Ruwe et al. 2016). Diese Umkehrung ist vermutlich auf die fragmentierten rRNA-Gene zurückzuführen. Mit Längen zwischen 20 und 200 nt bleiben die mitochondrialen rRNAs teilweise in der Größenselektion erhalten, während die apicoplastidäre rRNA keine wesentliche Rolle bei der Sequenzierung spielt, da diese durch die Größenselektion der cDNA-Bibliotheken weitestgehend ausgeschlossen werden.

### **3.2.3 CosRNAs im Apicoplasten liegen innerhalb von tRNA und rRNA Sequenzen**

Die bioinformatische Analyse der sequenzierten, dem Apicoplasten zugeordneten Fragmente zeigte, dass die potentiellen cosRNAs (*clustered organellar short* RNAs) ausnahmslos innerhalb von rRNA- und tRNA-Genen vorliegen. Es konnten auch in der enRNA keine cosRNAs außerhalb von diesen Regionen identifiziert werden. Dies ist überraschend, da die Plastiden von Pflanzen, wie in der Einleitung beschrieben eine Vielzahl dieser durch helikale Proteine generierten cosRNAs aufweisen. Aufgrund der insgesamt geringen Anzahl der zugeordneten Sequenzen kann die Existenz von cosRNAs in intergenischen oder kodierenden Regionen des Apicoplastengenoms nicht vollständig ausgeschlossen werden. Ob die Abwesenheit der cosRNAs biologische oder methodische Gründe hat, kann hier nicht aufgeklärt werden. Die apicoplastidären cosRNAs wurden nicht weiter behandelt.

### 3.2.4 Sequenzierung verbessert die Annotation der rRNA-Fragmente des Mitochondriums

Wie bereits erwähnt stammen 9 % bzw. 19 % der sequenzierten RNAs aus dem Mitochondrium. Eine Übersicht der Sequenzen, die dem mitochondrialen Genom zugeordnet werden konnten, sind in Abbildung 7 gezeigt. In dieser werden keine Einzelsequenzen dargestellt, sondern es wird die Anzahl der Sequenzen an einer bestimmten Position dargestellt, dies wird als Genomabdeckung bezeichnet. Um Bereiche mit geringerer Abdeckung hervorzuheben, wurde in der Abbildung 7 eine  $\log_2$  Skalierung gewählt. Obwohl die für die Sequenzierung verwendeten RNAs unterschiedlich hergestellt wurden, zeigen beide Datensätze eine sehr ähnliche Abdeckung des Genoms. Während die ribosomalen Fragmente durch eine hohe Abdeckung deutlich zu erkennen sind, zeigen die proteinkodierenden Gene eine geringere Abdeckung. Die im Folgenden beschriebenen cosRNAs und rRNA-Fragmente wurden, wenn nicht anders vermerkt, in beiden Datensätzen identifiziert.

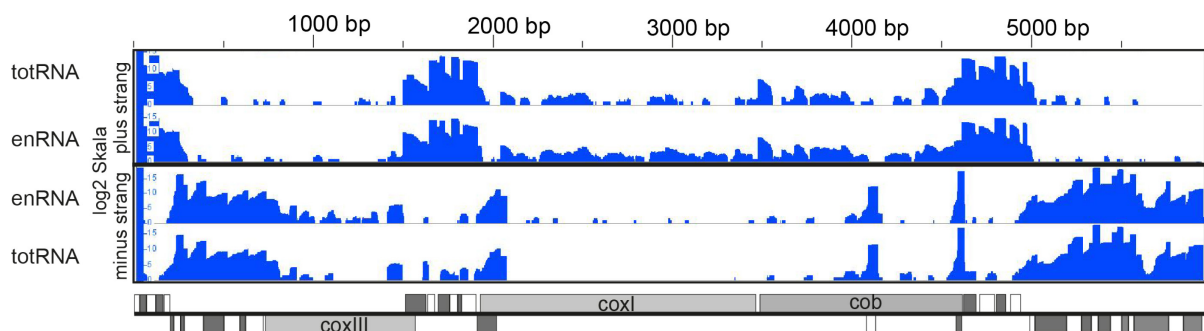


Abbildung 7: Genomkarte des Mitochondriums mit zugeordneten Sequenzen. Eine schematische Genomübersicht ist unter den zugeordneten *reads* gezeigt. Die strangspezifische Genomabdeckung zeigt die Häufigkeit der Sequenzen an dieser Genomposition. Die Sequenzierung sind strangspezifisch dargestellt. In der Genomkarte sind proteinkodierende Gene in Grau dargestellt, während dunkelgraue und weiße Elemente die ribosomalen Fragmente repräsentieren. Um schwächere Signale deutlicher zu machen wurde die Genomabdeckung (blau) mit einer  $\log_2$ -Skalierung dargestellt. Die Abdeckung ist in den Bereichen der ribosomalen Fragmente hoch, während die proteinkodierenden Gene eine geringere Abdeckung zeigen. Im Vergleich der beiden RNA-Präparationen zeigen sich trotz der unterschiedlichen Herstellungsmethoden sehr ähnliche Sequenzierungsergebnisse. Die detaillierte Beschreibung der sequenzierten Fragmente erfolgt im Text. (Hillebrand et al. 2018)

Insgesamt konnten von den 61 detektierten RNAs 52 in beiden Datensätzen gefunden werden, während 4 RNAs nur in der totRNA bzw. 5 nur in der enRNA-Sequenzierung identifiziert wurden.

Die meisten der von Feagin et al. 2012 beschriebenen rRNA-Fragmente konnten in der Analyse bestätigt werden. Auch einige der Fragmente die vorher nur durch seltene EST (*expressed sequence tags*) oder in *Northern*-Hybridisierungen detektiert wurden, konnten bestätigt werden. In Tabelle 2 werden die von Feagin et al. 2012 beschriebenen Fragmente mit den Ergebnissen dieser Sequenzierung verglichen. Nur für zwei der beschriebenen Fragmente konnten die Enden nicht eindeutig identifiziert werden bzw. nur in der enRNA (Tabelle 2, Fett geschrieben). An diesen Stellen ist die Genomabdeckung allerdings konstant hoch, wodurch die Bestimmung der Enden erschwert wird. Eine generelle Beobachtung ist, dass die 5'-Enden der Mehrzahl der sequenzierten Fragmente zumeist klar und scharf vom Hintergrund abgegrenzt sind, während die 3'-Enden wesentlich heterogener erscheinen. Einer der Gründe liegt in der unspezifische Polyadenylierungsaktivität für Transkripte im Mitochondrium. In dieser Sequenzierung kann nicht zwischen transkribierten und später angehängten Adeninen unterschieden werden (siehe Rehkopf et al. 2000). Für Transkripte mit einem oder mehreren Adeninen am 3'-Ende kann in dieser Arbeit also das 3'-Ende nicht genau definiert werden. In sieben Fällen zeigen Fragmente in der Sequenzierung ein 5'-Ende, das einige Basen stromauf bzw. stromabwärts der beschriebenen Position liegt (Tabelle 2, kursiv geschrieben). Fünf Fragmente, die vorher nur postuliert wurden, konnten ebenfalls bestätigt werden (siehe Tabelle 2 rRNA23t-rRNA27t). Die potentiell angehängten Adenine sind ebenfalls in Tabelle 2 vermerkt. Insgesamt konnte durch die Sequenzierung eine verbesserte Annotation des mitochondrialen Genoms erreicht werden.

## Ergebnisse

Tabelle 2: Vergleich der sequenzierten Fragmente mit den in Feagin et al. 2012 beschriebenen. **Fett** = nicht identifiziert, *kursiv* = unterschiedliches 5'-Ende, rRNA24-27t = postuliertes Fragment, n.i. = nicht identifiziert.

In Feagin et al. 2012 beschrieben		Eigene Sequenzierung		
Fragmentname	Genomposition (Strang)	Genomposition (Strang)	Adenine (3')	Positionsunterschiede
RNA16	3 - 33 (+)	3 – 34 (+)	1	
<b>RNA20</b>	34 - 71 (+)	n.i..	-	
RNA9	72 - 125 (+)	72-126 (+)	1-4	
<i>RNA17</i>	126 - 165 (+)	125-165 (+)	-	5' +1
RNA23t	171 - 203 (+)	166 – 205 (+)	5	5' +5
RNA23t	171 - 203 (+)	171 – 205 (+)	5	
RNA24t	262 - 224 (-)	262 – 223 (-)	6	
<b>LSUC</b>	225* - 204 (-)	n.i.	-	
RNA25t	283 - 262 (-)	283 – 263 (-)	-	
LSUG	389 - 283 (-)	389 – 289 (-)	-	
SSUB	505 - 390 (-)	505 – 405 (-)	-	
RNA1	593 - 506 (-)	593 – 506 (-)	3	
RNA15	624 - 594 (-)	624 – 593 (-)	4	
RNA10	724 - 625 (-)	724 – 624 (-)	4	
<i>LSUF</i>	1516 - 1630 (+)	1515 – 1578 (+)	-	5' +1
SSUE	1638 - 1680* (+)	1638 – 1680 (+)	1	
RNA2	1698 - 1763 (+)	1698 – 1725 (+)	-	
RNA26t	1764 - 1806 (+)	1764 – 1807 (+)	7	
RNA21	1807 - 1829 (+)	1807 – 1830 (+)	1	
RNA3	1830 - 1910 (+)	1830 – 1910 (+)	1	
SSUA	2023 - 1916 (-)	2030 – 1950 (-)	-	5' +7
RNA27t	4138 - 4085 (-)	4138 – 4084 (-)	2	
LSUB	4618 - 4586 (-)	4618 – 4584 (-)	2	
<i>RNA4</i>	4625 - 4696 (+)	4624 – 4698 (+)	2	5' +1
<i>RNA5</i>	4716 - 4802 (+)	4720 – 4803 (+)	6	5' -4
RNA6	4808 - 4865 (+)	4808 – 4868 (+)	6	
<i>RNA12</i>	4887 - 4945 (+)	4893 – 4944 (+)	1	5' -6
RNA18	4964 - 4988 (+)	4964 – 4988 (+)	-	
RNA13	5025 - 4996 (-)	5027 – 4997 (-)	-	
LSUA	5201 - 5026 (-)	5201 - ~5026 (-)	-	
RNA7	5283 - 5202 (-)	5283 – 5202 (-)	-	
RNA11	5340 - 5284 (-)	5340 – 5284 (-)	2	
RNA22	5378 - 5341 (-)	5379 – 5340 (-)	1	
SSUD	5446 - 5379 (-)	5446 -4378 (-)	2	
SSUF	5507 - 5447 (-)	5507 – 5446 (-)	6	
RNA14	5546 - 5508 (-)	5546 – 5503 (-)	5	
RNA19	5576 - 5547 (-)	5576 – 5546 (-)	3	
LSUE	5771 -5577 (-)	5771 – 5691 (-)	-	
LSUD	5854 - 5772 (-)	5854 – 5771 (-)	3	
<i>RNA8</i>	5954 - 5855 (-)	5953 - 5861 (-)	-	5' -1

### 3.2.5 Mitochondriale cosRNAs passen zu den 5'-Enden von beschriebenen rRNA-Fragmente

Eine faszinierende Beobachtung ergab sich bei der Analyse der Sequenzierdaten durch die in unserer Gruppe entwickelte Software *sRNA-Miner* (Ruwe et al. 2016). Diese wurde entwickelt, um potentielle PPR-Protein Bindestellen bzw. die zugehörige cosRNA zu identifizieren. Mit Parametern, die cosRNAs mit scharfen 5'-Enden und weniger definierten 3'-Enden identifizierten, konnten 25 cosRNAs (siehe Tabelle 3 und Tabelle 4) bestimmt werden.

Unter den durch den sRNA-Miner identifizierten RNAs befinden sich teilweise bereits beschriebene Transkripte. Dabei handelt es sich um kurze rRNA-Fragmente, die zwar als cosRNA identifiziert werden, jedoch mit Bezug auf die Transkriptenden und Genompositionen mit beschriebenen rRNA-Fragmente übereinstimmen. Diese wurden deshalb im Folgenden nicht weiter untersucht. Interessanterweise haben 12 der gefundenen cosRNAs ein 5'-Ende, welches deckungsgleich mit den 5'-Enden von rRNA-Genen ist. Das 3'-Ende stimmt nicht überein, sondern terminiert innerhalb des rRNA-Transkriptes.

Tabelle 3: Liste der gefundenen cosRNAs, die mit beschriebenen rRNA Fragmenten überlappen. \* = konnte nur in der enRNA Sequenzierung identifiziert werden.

cosRNA Name	rRNA Name	Genomposition (Strang)	Länge [nt]	Adenine (3')	Anzahl reads 5'-Ende	Position der cosRNA im Fragment
cosRNA1	LSUG	389 - 355 (-)	35	-	6812	5
cosRNA2	LSUG	389 - 333 (-)	57	-	6812	5
cosRNA3*	LSUF	1515 - 1554 (+)	40	-	370	5
cosRNA4	RNA2	1698 - 1725 (+)	28	-	13123	5
cosRNA5*	SSUA	2030 - 1972 (-)	59	-	1098	5
cosRNA6	RNA6	4849 - 4868 (+)	21	6	13016	3
cosRNA7	RNA12	4893 - 4944 (-)	52	-	2262	5
cosRNA8	LSUA	5201 - 5157 (-)	45	-	726	5
cosRNA9	RNA7	5283 - 5243 (-)	41	-	17260	5
cosRNA10	RNA7	5283 - 5226 (-)	58	-	17260	5
cosRNA11	RNA11	5340 - 5320 (-)	20	1	9989	5
cosRNA12	SSUF	5507 - 5478 (-)	30	-	158228	5
cosRNA13	LSUE	5771 - 5746 (-)	26	-	5216	5

Diese sind in Tabelle 3 mit den korrespondierenden rRNA-Genen aufgelistet. Von diesen konnten 2 (cosRNA3 und 5) nur in der enRNA-Sequenzierung identifiziert werden. In manchen Fällen wurden Fragmente mit dem gleichen 5'-Ende, aber unterschiedlichen 3'-Enden identifiziert (cosRNA1 und 2 bzw. RNA9 und 10). Diese Fragmente könnten auf eine schrittweise erfolgende Degradation der rRNA-Transkripte hindeuten, bei der stabilere Zwischenprodukte entstehen, die in der Sequenzierung identifiziert werden. Nur die cosRNA6 konnte dem 3'Ende eines Fragments zugeordnet werden.

### **3.2.6 Einige mitochondriale cosRNAs können in RNA-Hybridisierungen detektiert werden**

Um die Akkumulation ausgewählter cosRNAs mit einer alternativen Methode zu bestätigen, wurden 1,5 µg enRNA von *P. falciparum* parallel mit einem Größenstandard in einer RNA-Urea-PAGE aufgetrennt (Abbildung 8 B und D). Bei der Detektion der cosRNA12 wurde aus Verfügbarkeitsgründen pflanzliche gesamt RNA als Hybridisierungskontrolle genutzt; bei der Detektion der cosRNA11 wurde totRNA als Kontrolle verwendet (Abbildung 8 D). Die Detektion der spezifischen cosRNAs erfolgte mit am 5'-Ende radioaktiv markierten Oligonukleotiden, die komplementär zu den 5'-Enden der cosRNA sind.

Am 5'-Ende der rRNA11 konnte in der Sequenzierung die cosRNA11 identifiziert werden. Während die 5'-Enden dieser cosRNA deckungsgleich mit der reifen rRNA11 sind, endet die cosRNA11 nach 20 nt innerhalb der rRNA11 (Abbildung 8 A). Bei der cosRNA11 sind sowohl das 5'-Ende, als auch das 3'-Ende durch eine massive Änderung der Abdeckung scharf definiert. In der radioaktiven Detektion war in der aufgetrennten totRNA bei ca. 55 nt ein Signal erkennbar, welches mit der reifen rRNA11 korrespondiert (56 nt Länge). Für die cosRNA11 (20 nt Länge) war kein entsprechendes Signal vorhanden (siehe Abbildung 8 B, Spur 1). Zusätzlich zu der reifen rRNA11 ist ein Fragment von mehr als 70 nt erkennbar. Es könnte sich hierbei um einen polycistronischen Vorläufer handeln, bei dem die rRNA11 entweder mit dem stromauf liegenden rRNA-Gen (rRNA22, 38 nt Länge), oder mit der stromab liegenden rRNA7 (82 nt Länge) in einer RNA vorliegt. In der enRNA war neben den gleichen Signalen wie in der gesamt RNA, ein Fragment in der Größe der cosRNA11 sichtbar. Zwischen diesem und der reifen rRNA11 sind

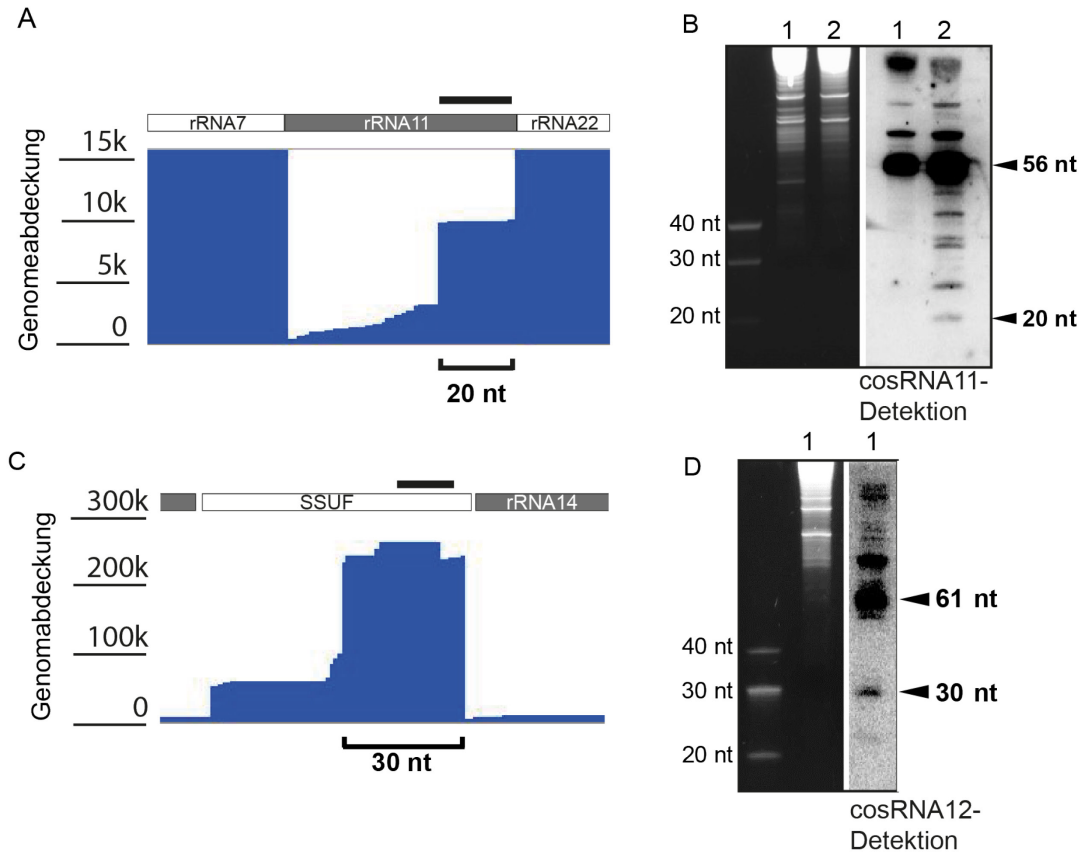


Abbildung 8: Nachweis ausgewählter cosRNAs durch radioaktive Detektion. **A)** In der rRNA11 wurde ein Fragment von 20 nt sequenziert (cosRNA11). Während das scharfe 5'-Ende mit dem 5'-Ende der rRNA11 korrespondiert, liegt das 3'-Ende der cosRNA11 innerhalb der rRNA11. **B)** Die aufgetrennte gesamte RNA (Spur 1) zeigte in der Detektion kein Signal für die cosRNA11, jedoch ist das reife Fragment mit 56 nt erkennbar. Bei den größeren Banden könnte es sich um einen polycistronischen Vorläufer aus der rRNA11 mit entweder rRNA7 oder rRNA22 handeln. In der enRNA (Spur 2) kann ein Signal in der Größe der cosRNA11 identifiziert werden. Zwischen der reifen rRNA11 und der cosRNA11 werden zusätzliche Signale detektiert. **C)** Die Sequenzierung zeigte innerhalb der rRNA SSUF ein 30 nt langes Fragment mit scharfem 5'- und 3'-Ende. Hier ist nur die Genomabdeckung der enRNA gezeigt, die totRNA zeigt ein äquivalentes Bild. **D)** Bei der Detektion der markierten Sonde wurden in der enRNA (Spur 1) Signale bei 30 nt und bei 61 nt detektiert, dies entspricht der Größe der cosRNA12 bzw. der reifen SSUF rRNA. Oberhalb der reifen SSUF rRNA konnten einige größere Fragmente detektiert werden, hierbei kann es sich um polycistronische Vorläufer handeln. Als Kontrolle wurde in diesem Fall pflanzlichen gesamte RNA verwendet, die keine Signale lieferte (nicht gezeigt). Die für die Detektion verwendeten Oligonukleotide sind als schwarzer Balken eingezeichnet. (Hillebrand et al. 2018)

weitere Signale erkennbar. Diese können durch die schrittweise Degradation der reifen rRNA11 entstehen und beispielsweise durch Stellen, an denen Endo- bzw.



Exonukleasen stoppen, definiert werden. Zwar sind diese Fragmente nicht in der Sequenzierung identifizierbar, allerdings sind Fragmente mit identischem 5'-Ende und variablem 3'-Ende erkennbar.

Innerhalb der rRNA SSUF konnte in der Sequenzierung eine cosRNA am 5'-Ende identifiziert werden (siehe Abbildung 8 C). Diese cosRNA12 war besonders dominant in der Sequenzierung mit etwa 160.000 *reads* (totRNA). In der Hybridisierungskontrolle der pflanzlichen Kontroll-RNA wurde kein Signal detektiert. In der enRNA konnte jedoch sowohl für die reife SSUF rRNA (61 nt) als auch für die cosRNA12 (30 nt) jeweils ein klares Signal identifiziert werden (siehe Abbildung 8 D). Oberhalb der SSUF rRNA sind einige zusätzliche Signale erkennbar. Dabei könnte es sich beispielsweise um polycistronische Vorläufer-Transkripte handeln, die neben der SSUF rRNA weitere Transkripte enthalten. Im Gegensatz zur cosRNA11 konnten keine Degradationsprodukte zwischen dem reifen Fragment und der cosRNA12 detektiert werden. Dies könnte darauf hindeuten, dass bei der Degradation der SSUF rRNA nur die cosRNA12 als stabiles Zwischenprodukt entsteht. Zusammenfassend konnten zwei cosRNAs in der enRNA mittels *Northern*-Hybridisierung detektiert und damit die Ergebnisse der Sequenzierung an dieser Stelle bestätigt werden.

### **3.2.7 Einige mitochondriale cosRNAs definieren die Transkriptenden von mRNAs, andere liegen in intergenischen Bereichen**

Während die oben beschriebenen cosRNAs (siehe 3.2.5) spezifischen Transkripten zugeordnet werden können, wurden auch cosRNAs identifiziert, bei denen dies nicht möglich war. Diese sind in Tabelle 4 aufgelistet (cosRNA14-25). Sie unterscheiden sich nicht nur in der Genomposition, sondern auch in der Häufigkeit mit der sie sequenziert wurden. In beiden Datensätzen zeigen sie meist eine niedrigere Abdeckung als die cosRNAs1-13.

Fünf der cosRNAs (cosRNAs14, 16, 18, 19, 21) liegen direkt stromauf des 5'-Endes von rRNA-Genen. Diese cosRNAs könnten durch die Prozessierung des reifen Fragments aus einem polycistronischen Transkript entstehen. Bei diesem Schritt könnte mit der Prozessierung des rRNA-Fragments auch die jeweilige

cosRNA freigesetzt werden. Als Beispiel ist in Abbildung 9 A cosRNA16 gezeigt. Diese liegt unmittelbar stromauf der rRNA LSUF.

Tabelle 4: cosRNAs die außerhalb von rRNA-Genen liegen. \* = nicht in enRNA-Sequenzierung identifiziert, # = nur in enRNA-Sequenzierung identifiziert.

cosRNA Name	Genomposition (Strang)	Länge [nt]	Adenine (3')	Anzahl reads am 5'-Ende	Anmerkung
cosRNA14	166 - 205 (+)	40	5	550	vor rRNA Fragment
cosRNA15	787 - 725 (-)	63	-	68	3'-Ende der <i>coxIII</i> mRNA
cosRNA16	1492 - 1513 (+)	22	-	145	vor rRNA Fragment
cosRNA17	1630 - 1601 (-)	30	-	113	intergenische <i>coxIII</i> <> SSUA
cosRNA18	1638 - 1725 (+)	88	-	n.d.	vor rRNA Fragment
cosRNA19	2067 - 2031 (-)	37	-	230	vor rRNA Fragment
cosRNA20	3480 - 3537 (+)	58	-	153	5'-Ende der <i>cob</i> mRNA
cosRNA21*	4866 - 4944 (+)	79	-	~120	vor rRNA Fragment
cosRNA22#	203 - 238 (+)	36	-	1204	3' der rRNA23t
cosRNA23#	893 - 855 (-)	39	-	60	kodierenden Region <i>coxIII</i> 3'
cosRNA24#	1486 - 1412 (-)	75	-	70	5' der <i>coxIII</i> mRNA
cosRNA25#	4697 - 4719 (+)	26	3	671	überlappend rRNA4 3' und 5'rRNA5

Drei der cosRNAs konnten den mRNAs des mitochondrialen Genoms zugeordnet werden. Alle im weiteren hier herangezogenen Transkriptenden der mRNAs wurden in Rehkopf et al. 2000 bestimmt. Am 3'-Ende der *coxIII* mRNA liegt die cosRNA15. Das 5'-Ende der cosRNA20 befindet sich im nicht kodierenden Bereich zwischen *coxI* und *cob* und ist deckungsgleich mit dem beschriebenen 5'-Ende der *cob* mRNA. Das 3'-Ende ist kaum differenziert und liegt innerhalb der kodierenden Sequenz der *cob* mRNA (siehe Abbildung 9 B). Als drittes konnte die cosRNA24 in der enRNA-Sequenzierung identifiziert werden. Diese ist passend zum beschriebenen 5'-Ende der *coxIII* mRNA.

Die cosRNA17 liegt im intergenischen Bereich zwischen dem 5'-Ende der *coxIII* mRNA und dem stromauf liegenden rRNA SSUA-Gen. Damit stellt sie die einzige cosRNA dar, die keinem direkten Transkript zugeordnet werden kann. Möglicherweise ist diese cosRNA ein stabiler Teil eines polycistronischen Vorläufertranskripts, das SSUA und *coxIII* enthält. Die cosRNA17 ist auch die einzige identifizierte RNA, die ähnlich wie pflanzliche cosRNAs im intergenischen Bereich liegt. Die Analyse der cosRNA20 und cosRNA17 in Northern-

Hybridisierungen blieb erfolglos, vermutlich aufgrund des niedrigen

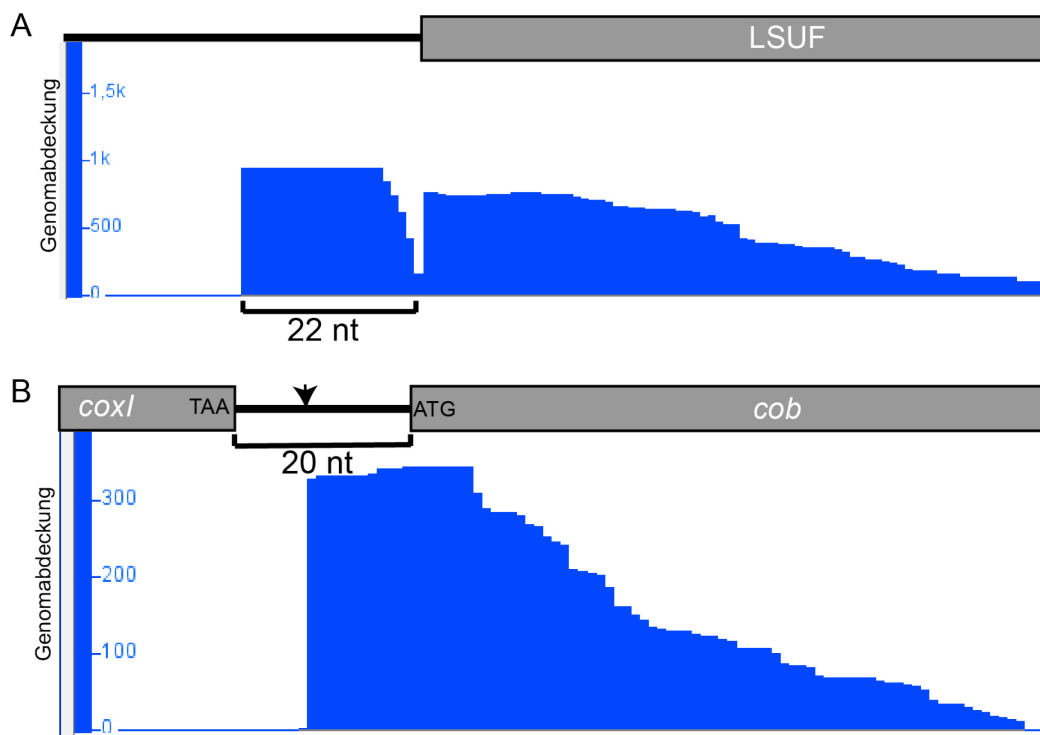


Abbildung 9: Beispiele für cosRNAs außerhalb von rRNA Fragmenten. **A)** Unmittelbar am 5'-Ende der rRNA LSUF konnte die cosRNA16 (22 nt Länge) detektiert werden, die vermutlich Teil eines Vorläufermoleküls ist. Sowohl 5'- als auch 3'-Ende der cosRNA sind klar definiert. Das 3'-Ende zeigt einen deutlichen Abfall der Genomabdeckung zwischen LSUF und cosRNA16. **B)** Die cosRNA 20 liegt in der 20 nt langen, nicht-kodierenden Region zwischen den mRNAs *coxI* und *cob*. Hier ist das 5'-Ende scharf abgegrenzt und passt zum 5'-Ende der monocistronischen *cob* mRNA (Pfeil, Rehkopf et al. 2000). Das 3'-Ende der cosRNA ist kaum definiert und liegt innerhalb der kodierenden Sequenz der *cob* mRNA. Beide gezeigten cosRNAs konnten in Northern-Detektionen nicht identifiziert werden.

Expressionsniveaus dieser cosRNAs (siehe Tabelle 4). Zusammenfassend kann man sagen, dass in den Mitochondrien von *P. falciparum* cosRNAs identifiziert werden konnten. Diese unterscheiden sich allerdings von den cosRNAs pflanzlicher Organellen, da diese cosRNAs meist in intergenischen Bereichen vorliegen, während die cosRNAs im Mitochondrium von *P. falciparum* größtenteils innerhalb von definierten Transkripten identifiziert wurden.

### 3.3 Identifikation einer neuen Familie von helikalen Proteinen in *Plasmodium*

Die Identifikation der cosRNAs in den Mitochondrien von *P. falciparum* führte zu der Hypothese, dass helikale Proteine wie PPR und OPR-Proteine an der Entstehung dieser cosRNAs beteiligt sein könnten. Um diese Frage zu untersuchen, wurde eine genomweite Suche nach helikalen Proteinen auf das Proteom von *P. falciparum* durchgeführt. Die Suchmotive für die repetitiven Elemente wurden mit Hilfe des MEME (**M**ultiple **E**m for **M**otif **E**licitation) Algorithmus erstellt. Die Analyse des Proteoms von *Plasmodium* mit den erstellten Suchmotiven wurde mittels MAST (**M**otif **A**lignment & **S**earch **T**ool) durchgeführt. Sowohl MEME, als auch MAST sind Teil des *MEME-Suite* Software Pakets<sup>6</sup>. Aufgrund ihrer in der Einleitung beschriebenen Bedeutung im RNA Metabolismus von pflanzlichen Organellen, wurde die Suche auf PPR- und OPR-Proteinmotive fokussiert. Bei der Suche mit den erstellten Suchmotiven im Proteom von *P. falciparum* (Proteindaten von der PlasmoDB, Aurrecoechea et al. 2009) im MAST-Algorithmus wurden unter den gefundenen Proteinen nur solche verwendet, die mindestens zwei Motivwiederholungen zeigen (maximaler Abstand 3 Aminosäuren zwischen den beiden Wiederholungen) und einen e-Wert von <0,0001 hatten. Der e-Wert gibt bei den Suchen eine Wahrscheinlichkeit an, mit der das gefundene Motiv auch in einem zufällig generierten Proteindatensatz der gleichen Größe gefunden werden könnte. Hohe E-Werte deuten dabei auf ein zufälliges Motiv hin, während niedrige e-Werte auf signifikante Treffer hinweisen.

#### 3.3.1 Genomweite Suche mit einem PPR-Konsensus Motiv identifiziert keine neuen Proteine

Das PPR-Suchmotiv wurde mit ausgewählten, beschriebenen PPR-Proteinen aus *A. thaliana* und *Z. mays* erstellt. Eine Liste der verwendeten PPR-Proteine ist im Anhang aufgeführt (Anhang Tabelle 4). Dabei wurde der MEME-Algorithmus auf multiple Wiederholungen in einem Protein und auf eine definierte Länge von 35 Aminosäuren eingestellt. Die anschließend durchgeführte MAST-

---

<sup>6</sup> <http://meme-suite.org/index.html>

Suche auf das Proteom von *P. falciparum* lieferte zwei Proteine mit multiplen PPR-Motiven. Diese beiden Proteine wurden bereits beschrieben und als PPR Proteine klassifiziert (Hicks et al. 2018).

### **3.3.2 Suche nach OPR-ähnlichen Proteinen identifiziert ein neues repetitives Element in *P. falciparum* Proteinen**

Für die Suche nach OPR-ähnlichen Proteinen wurde das Suchmotiv aus den 42 OPR-Proteinen gebildet, die in Rahire et al. 2012 beschreiben sind. Die Proteinsequenzen wurden dazu aus der Datenbank des JGI<sup>7</sup> (Joint Genome Institute) heruntergeladen (Nordberg et al. 2014). Die Erstellung des Suchmotivs im MEME-Algorithmus wurde auf 38 Aminosäuren lange Motive beschränkt. Gleichzeitig wurde nach multiplen Wiederholungen des Motivs in Proteinen gesucht. Das so generierte Suchmotiv war weitestgehend identisch mit dem in Rahire et al. 2012 verwendeten Motiv (siehe Abbildung 10). Zur Überprüfung des Suchmotivs wurde eine reziproke Suche im Proteom von *C. reinhardtii* durchgeführt. Dabei wurden 46 OPR-Proteine nach den definierten Kriterien identifiziert. Bei der anschließenden Suche im Proteom von *P. falciparum* wurden nur drei Proteine mit mehr als zwei OPR-Motiven identifiziert, wobei keines unterhalb der definierten e-Wertgrenze lag.

Da sowohl OPR-, als auch PPR-Proteine einen Längenpolymorphismus besitzen, wurde die Suche mit verkürzten 36- bzw. 37-AS langen Suchmotiven wiederholt. Dabei wurden die gleichen Ausgangsproteine wie oben beschrieben genutzt. Hier konnten sechs Proteine mit mindestens 2 Wiederholungen in direkter Nähe identifiziert werden. Diese Proteine waren für beide Suchmotive identisch, mit dem Unterschied, dass bei dem 36-AS langen Suchmotiv eine Lücke von einer zusätzlichen Aminosäure zwischen den Motiven eingefügt wurde. Aus diesem Grund wurde nur mit dem 37-AS langen Motiv weitergearbeitet. Die neu identifizierten Proteine wurden als Grundlage genutzt um ein neues, *P. falciparum* spezifisches Proteinmotiv zu generieren. Da es sich nur um wenige Proteine handelte, wurden die OPR-Proteine von *C. reinhardtii* nicht in die Erstellung des

---

<sup>7</sup> <https://genome.jgi.doe.gov>

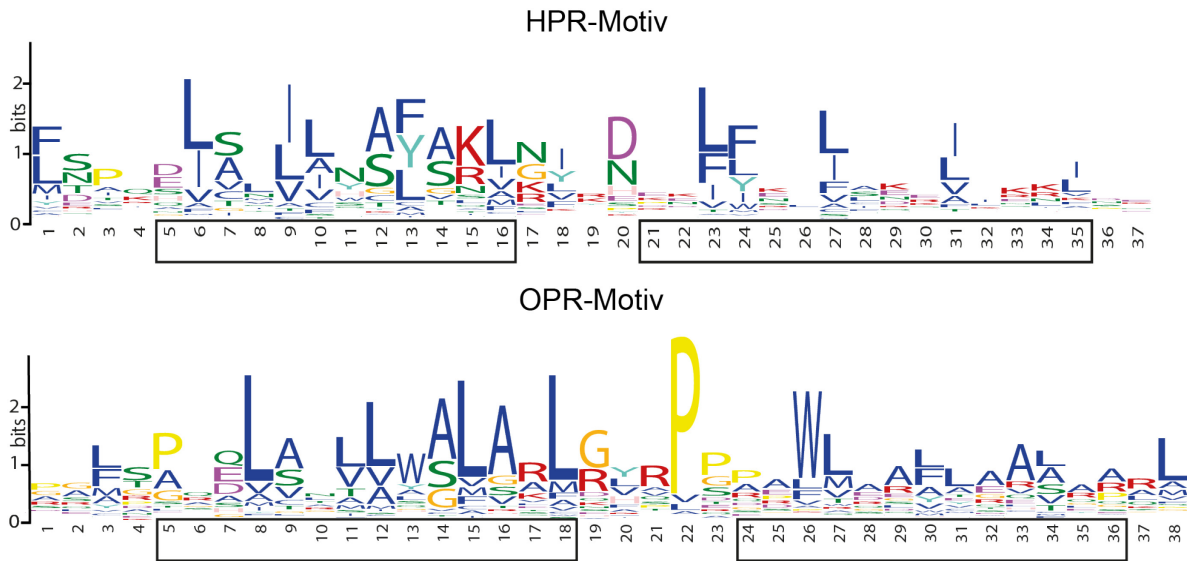


Abbildung 10: Sequenzlogo des *Heptatricopeptide-repeat* Motivs (HPR). Im OPR-Motiv (unten) ist die PPPEW-Domäne (OPR-Motiv Position 22-26) klar zu erkennen, diese Domäne fehlt im HPR Motiv (oben). Beide Motive formen eine Struktur aus zwei antiparallelen Helices (schwarze Kästen) unterbrochen von einem unstrukturierten Bereich. Weiterhin zeigen beide Motive gerade im ersten der zwei helikalen Elemente konservierte Leucine. Die Aminosäuren in den Positionen 1-14 des HPR-Motivs zeigen eine deutliche Konservierung mit den AS der Positionen 3-16 des OPR Motivs. Diese Konservierung ist in der zweiten Hälfte des Motivs deutlich reduziert. Grundsätzlich scheint die zweite Helix beider Motive bis auf einige AS schwächer konserviert. (Hillebrand et al. 2018, Bailey et al. 2009)

Suchmotiv eingeschlossen, um eine Überlagerung des *Plasmodium*-spezifischen Motivs zu verhindern. Mit dem neu erstellten Motiv konnten in einer weiteren MAST-Suche zusätzliche Proteine im Proteom von *P. falciparum* identifiziert werden. Durch mehrere Iterationen der MAST-Suche wurde das Motiv weiter verfeinert. Neu gefundene Proteine wurden nach jeder Suche zu den Proteinen hinzugefügt und ein neues Suchmotiv erstellt. Nach 4 Durchläufen wurden keine weiteren Proteine identifiziert. Die 22 identifizierten Proteine sind in Abbildung 11 dargestellt, eine genauere Beschreibung erfolgt in 3.3.3.

Mit dem finalen Suchmotiv aus *P. falciparum* wurde auch in den verwandten Organismen *P. berghei* und *T. gondii* nach ähnlichen Proteinen gesucht. Auch hier konnten mehrere Proteine mit multiplen Motivwiederholungen identifiziert werden (siehe Abbildung 12 und Abbildung 13). Um ein für Apicomplexa spezifisches Suchmodell zu generieren, wurden insgesamt 299 Motive aus allen drei Spezies

kombiniert und ein finales Suchmotiv generiert (siehe Abbildung 10). Im Vergleich mit dem OPR-Motiv fällt auf, dass neben dem Längenunterschied des neuen Motivs auch die in OPR-Motiven konservierte PPPEW-Domäne fehlt. Aufgrund des Längenunterschiedes und der veränderten Sequenz definiert das neue Motiv vermutlich eine neue Proteinfamilie mit repetitiven helikalen Motiven, die wir aufgrund der 37 Aminosäuren langen Elemente und ihrer Verwandtschaft zu OPR-Proteinen *Heptatricopeptide-repeat* Proteine (HPR) genannt haben. Wie bereits PPR- und OPR-Proteine zeigt auch das HPR-Motiv eine Struktur aus zwei antiparallelen Helices und einem kleinen, unstrukturierten Bereich dazwischen (Abbildung 10, schwarze Kästen).

### **3.3.3 Die Analyse der HPR-Proteine in *P. falciparum* deutet auf eine fast exklusiv mitochondriale Lokalisation hin**

Insgesamt wurden im Genom von *P. falciparum* 22 Proteine identifiziert, die mindestens zwei Motivwiederholungen und einen e-Wert von  $<0,0001$  besitzen. Diese Proteine sind schematisch in Abbildung 11 dargestellt. Eine tiefergehende Analyse der Proteine zeigte, dass fünf dieser Proteine zusätzlich zu den HPR-Motiven eine C-terminale RAP-Domäne besitzen. Diese potentielle Nuklease-Domäne wurde ebenfalls für einige OPR-Proteine beschrieben und könnte auf eine Funktion der HPR-Proteine im RNA-Metabolismus hindeuten (Eberhard et al. 2011, Rahire et al. 2012). Um potentielle weitere Proteindomänen zu identifizieren, wurden die Proteine mit der Proteindatenbank PFAM<sup>8</sup> untersucht. Außer der RAP-Domäne wurden keine weiteren konservierten Domänen identifiziert. Jedoch ist nicht ausgeschlossen, dass individuelle Domänen auf den Proteinen liegen, die nicht in der PFAM-Datenbank enthalten sind. Die für einen Teil der OPR-Proteine aus *C. reinhardtii* identifizierten FAST-Domänen konnten nicht identifiziert werden (Eberhard et al. 2011, Jourdain et al. 2017). Die HPR-Motive liegen meist in dichter Folge hintereinander auf einem Protein (siehe Abbildung 11). Diese Akkumulationen der HPR-Motive werden an manchen Stellen durch Bereiche unterbrochen, in denen keine HPR-Motive identifiziert werden (siehe Abbildung 11, z.B. PF3D7\_1221500). Durch diese Unterbrechungen werden die HPR-Wiederholungen

---

<sup>8</sup> <https://pfam.xfam.org/>

## Ergebnisse

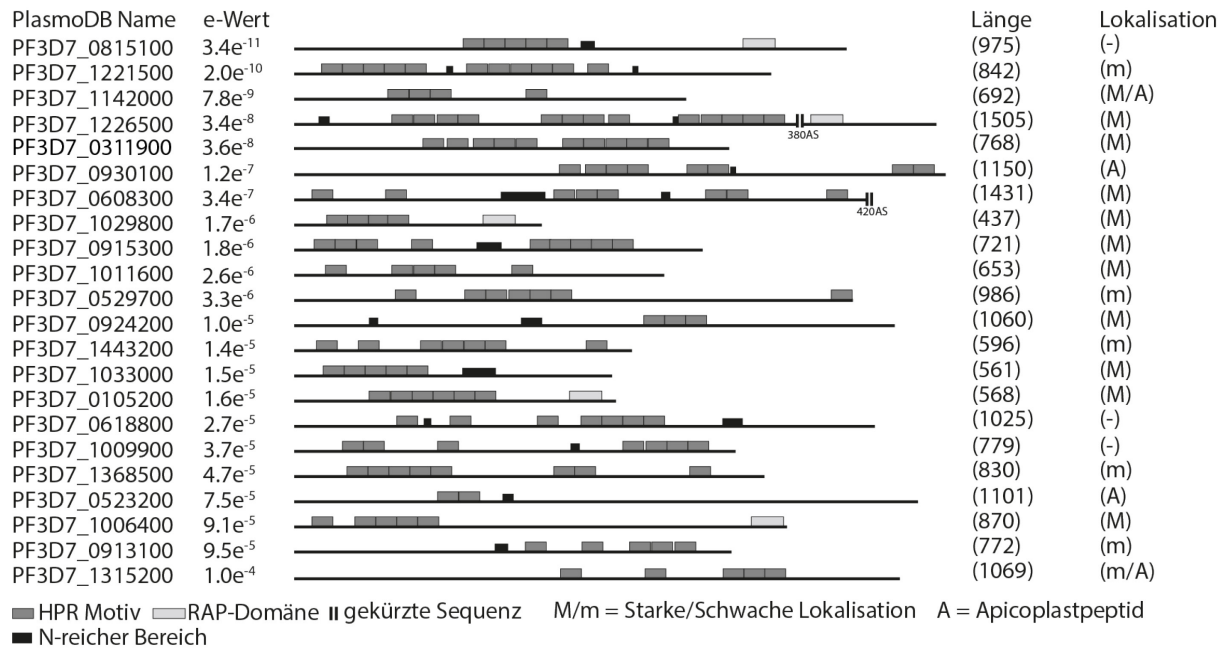


Abbildung 11: Schematische Darstellung der identifizierten HPR-Proteine in *P. falciparum*.

Die durch den MAST-Algorithmus identifizierten Proteine zeigen mindestens zwei Wiederholungen des HPR-Motivs (dunkel graue Blöcke). Interessanterweise besitzen fünf dieser Proteine eine C-terminale RAP-Domäne (hell graue Blöcke). Die für *P. falciparum* beschriebenen Asparagin-reichen Abschnitte wurden ebenfalls identifiziert und markiert (schwarze Blöcke, Muralidharan und Goldberg 2013). Die Vorhersage der Lokalisation deutet bei fast allen Proteinen auf eine Lokalisation im Mitochondrium hin. M = starke Lokalisationsvorhersage >0.5, m= schwache Lokalisationsvorhersage > 0,3 und <0,5. Nur vier Proteine sind potentiell im Apicoplasten (A) lokalisiert; bei zweien davon besteht die Möglichkeit einer Doppellokalisierung in beide Organellen. (Hillebrand et al. 2018)

in mehrere Abschnitte unterteilt. Die Unterbrechung zwischen diesen HPR-Motivakkumulationen haben teilweise eine Länge von 37 AS und formen in Sekundärstrukturvorhersagen ebenfalls helikale Elemente. Bei diesen Lücken handelt es sich vermutlich um stärker degenerierte HPR-Motive, die vom Suchmotiv nicht identifiziert werden. Zwischen einigen HPR-Motiven befinden sich 1-10 zusätzliche AS. Dies könnte auf einen Längenpolymorphismus der HPR-Motive hindeuten. Etwa 30 % der gesamten Proteine von *P. falciparum* besitzen Asparagin-reiche Bereiche deren Funktion unbekannt ist (Muralidharan und Goldberg 2013). Die identifizierten Proteine wurden ebenfalls nach diesen Bereichen identifiziert (Abbildung 11, schwarze Blöcke). Es wurden einige Asparagin-reiche Domänen in den HPR-Proteinen gefunden, diese überlappen nicht mit den HPR-Bereichen. Sie



sind teilweise zwischen HPR-Motiven zu finden (siehe Abbildung 11, bspw. PF3D7\_1221500, PF3D7\_0915300).

Wie in der Einleitung beschrieben, lokalisieren die meisten PPR- und OPR-Proteine in die Plastiden bzw. Mitochondrien. Um eine potentielle Organellenlokalisierung für die HPR-Proteine zu untersuchen wurde eine bioinformatische Vorhersage der Lokalisation mit verschiedenen Algorithmen durchgeführt. Für die mitochondriale Lokalisationswahrscheinlichkeit wurde TargetP 1.1 (Emanuelsson et al. 2000) und MitoProt II (Claros und Vincens 1996) verwendet. Für die plastidäre Lokalisationswahrscheinlichkeit wurden PATS (*Prediction of apicoplast-Targeted Sequences*) (Zuegge et al. 2001) und PlasmoAP der PlasmoDB (Foth et al. 2003) verwendet. Die meisten HPR-Proteine zeigen bei der Lokalisationsvorhersage eine hohe Wahrscheinlichkeit für eine mitochondriale Lokalisation. Nur in 4 Proteinen wurde ein potentielles Apicoplasten-Translokationspeptid identifiziert. Von diesen 4 Proteinen besitzen 2 gleichzeitig eine hohe mitochondriale Lokalisationswahrscheinlichkeit. Dies könnten auf eine potentielle, duale Lokalisation in beiden Organellen hinweisen, wurde jedoch nicht weiter analysiert. Da cosRNAs nur in den Mitochondrien identifiziert werden konnten, erfolgten keine weiteren Untersuchungen dieser 4 Proteine. Es wäre jedoch eine spannende zukünftige Untersuchung, da die Mechanismen und Faktoren der apicoplastidären Expression größtenteils unbekannt sind (siehe 1.3.2).

### **3.3.4 HPR-Proteine konnten auch in *P. berghei* identifiziert werden**

Wie bereits erwähnt, konnten auch in anderen apicomplexen Parasiten HPR-Proteine identifiziert werden. Von besonderem Interesse für die folgenden Experimente war dabei der mit *P. falciparum* nah verwandte Parasit *P. berghei*. Der für Menschen ungefährliche Erreger verursacht Malaria in Nagetieren, wie etwa Mäusen. In *P. berghei* wurden 22 HPR-Proteine identifiziert (siehe Abbildung 12). Von diesen Proteinen enthalten drei Proteine eine C-terminale RAP-Domäne. Die wie oben beschrieben durchgeführte Lokalisationsvorhersage (siehe 3.3.3) für diese Proteine deutet ebenfalls auf eine mitochondriale Lokalisation hin. Sechzehn der detektierten Proteine sind Orthologe zu den in *P. falciparum* identifizierten HPR-Proteinen (siehe Tabelle 5).

## Ergebnisse

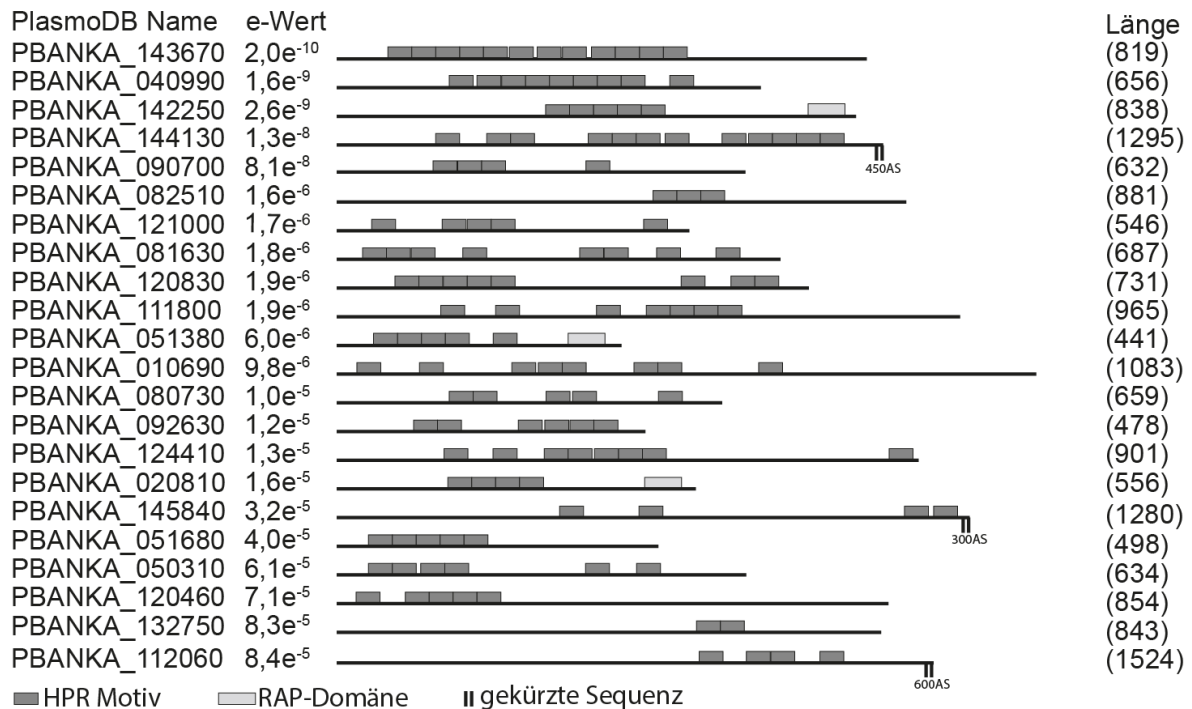


Abbildung 12: HPR-Proteine in *P. berghei*. Die durch den MAST-Algorithmus identifizierten Proteine zeigen mindestens zwei Wiederholungen des HPR-Motivs (dunkel graue Blöcke). Auch hier besitzen drei Proteine eine C-terminale RAP-Domäne (hell graue Blöcke). Für die Darstellung wurden 3 Proteine gekürzt dargestellt (Doppelstriche). Auch bei den in *P. berghei* identifizierten HPR-Proteinen deutet die Lokalisationsvorhersage auf eine überwiegend mitochondriale Lokalisation hin (Daten nicht gezeigt).

Für 6 Proteine in *P. berghei* konnte das Ortholog in *P. falciparum* nicht identifiziert werden (siehe Tabelle 5, Fett geschrieben). Gleichzeitig wurden in *P. falciparum* 6 HPR-Proteine identifiziert, für die das orthologe HPR-Protein in *P. berghei* nicht detektiert wurde (siehe Tabelle 5, Fett geschrieben). Die Zuordnung der Orthologe basiert auf den Daten der PlasmoDB. Die meisten der fehlenden Proteine wurden in der Suche aufgrund ihrer e-Werte ausgeschlossen. Es ist zu vermuten, dass es sich bei diesen ebenfalls um HPR-Proteine handelt, jedoch mit stärker degenerierten HPR-Motiven die zu einem höheren e-Wert führen.

Tabelle 5: Orthologe von HPR-Proteinen in *P. berghei* und *P. falciparum*. Bei fett geschriebene Proteine wurden in der MAST-Suche nicht identifiziert.

HPR-Protein in <i>P. berghei</i>	Ortholog in <i>P. falciparum</i>
PBANKA_143670	PF3D7_1221500
PBANKA_040990	PF3D7_0311900
PBANKA_142250	PF3D7_0815100
PBANKA_144130	PF3D7_1226500
PBANKA_090700	PF3D7_1142000
PBANKA_082510	PF3D7_0924200
PBANKA_121000	PF3D7_1011600
PBANKA_081630	PF3D7_0915300
PBANKA_120830	PF3D7_1009900
PBANKA_111800	PF3D7_0618800
PBANKA_051380	PF3D7_1029800
PBANKA_010690	PF3D7_0608300
PBANKA_080730	<b>PF3D7_0317900</b>
PBANKA_092630	<b>PF3D7_1122000</b>
PBANKA_124410	PF3D7_0529700
PBANKA_020810	PF3D7_0105200
PBANKA_145840	<b>PF3D7_1245200</b>
PBANKA_051680	PF3D7_1033000
PBANKA_050310	<b>PF3D7_1018900</b>
PBANKA_120460	PF3D7_1006400
PBANKA_132750	<b>PF3D7_1464100</b>
PBANKA_112060	<b>PF3D7_0621700</b>
<b>PBANKA_083080</b>	PF3D7_0930100
<b>PBANKA_130710</b>	PF3D7_1443200
<b>PBANKA_114440</b>	PF3D7_1368500
<b>PBANKA_123800</b>	PF3D7_0523200
<b>PBANKA_081410</b>	PF3D7_0913100
<b>PBANKA_141370</b>	PF3D7_1315200

### 3.4 HPR-Proteine außerhalb von *Plasmodium*

PPR-Proteine sind in vielen Eukaryoten zu finden, konnten aber besonders zahlreich in Pflanzen identifiziert werden (siehe 1.4.1). Um zu untersuchen ob HPR-Proteine eine vergleichbare Verteilung haben, wurde mit dem HPR-Motiv in repräsentativen Vertretern verschiedener phylogenetischer Gruppen nach HPR-Proteinen gesucht. Parallel wurde die gleiche Suche mit dem OPR-Motiv durchgeführt. Während dieser Suchen mit beiden Suchmotiven wurden keine

anderen vergleichbaren helikalen Proteine (bspw. PPR-Proteine, TPR-Proteinen, HAT-Proteine) identifiziert. Damit kann angenommen werden, dass die Suchmotive spezifisch für ihre jeweiligen Motive sind. In einigen Organismen wurde eine Überlappung zwischen OPR- und HPR-Proteinen festgestellt, bei dem ein Protein mit beiden Motiven gefunden werden konnte. Die Ergebnisse dieser Suchen sind in Abbildung 13 dargestellt. Die Darstellung der Suchergebnisse erinnert zwar an einen phylogenetischen Baum, jedoch wurden keine Berechnungen der evolutionären Distanzen durchgeführt. Es soll lediglich die Zugehörigkeit der Organismen zu bestimmten Taxa deutlich gemacht werden. Die Länge der Äste ist kein Indikator für die evolutionäre Distanz.

### 3.4.1 HPR-Proteine werden in fast allen untersuchten Alveolata identifiziert

Neben *P. berghei* und *P. falciparum* wurde in *T. gondii*, einem mit *Plasmodium* verwandten Parasiten, eine vergleichbare Anzahl von HPR-Proteinen identifiziert. In *Cryptosporidium parvum* und *Cryptosporidium hominis* konnten keine HPR-Proteine identifiziert werden. Dies ist besonders faszinierend, da die Gattung *Cryptosporidium* kein mitochondriales Genom besitzt und das mitochondriale Kompartiment der Zellen ausschließlich von kernkodierten Proteinen abhängig ist (Henriquez et al. 2005). Dies ist ein weiterer Hinweis darauf, dass HPR-Proteine potentiell mit der Expression des mitochondrialen Genoms im Zusammenhang stehen.

Zusammen mit den Apicomplexa werden auch photoautotrophe Organismen wie Chromerida in den Alveolata zusammengefasst. Bei den Chromerida handelt es sich um einzellige, photosynthetische Organismen, die im maritimen Lebensraum vorkommen. In zwei Vertretern der Chromerida, *Chromera velia* und *Vitrella brassicaformis* wurden 58 bzw. 70 HPR-Proteine identifiziert. In den mitochondrialen Genomen der beiden Chromerida wurden fragmentierte rRNA-Gene und einige wenige proteinkodierende Gene nachgewiesen (Flegontov et al. 2015). Das mitochondriale Genom des Dinoflagellaten *Symbiodinium minutum* enthält ebenfalls fragmentierte rRNA-Gene und es wurden deutliche Parallelen zwischen dem Gengehalt der mitochondrialen Genomen von *Plasmodium* und *S. minutum* nachgewiesen (Shoguchi et al. 2015). In *S. minutum* konnten 73 HPR-Proteine und nur 3 OPR-Proteine; in *S. microadriaticum* 49 HPR Proteine und 2

## Ergebnisse

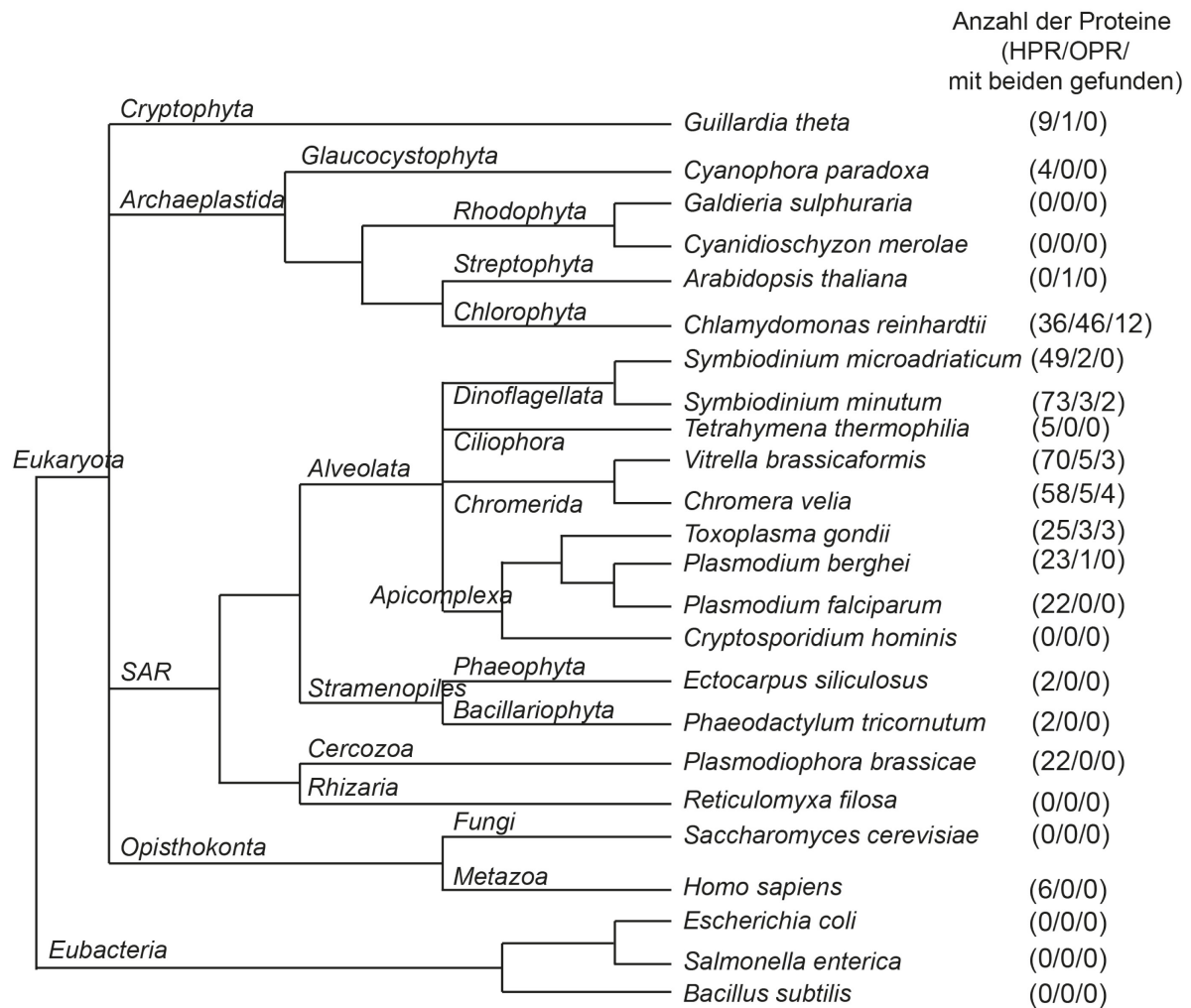


Abbildung 13: Verteilung der HPR-Proteine in verschiedenen Spezies. In den Proteomen verschiedener Organismen wurde mit einer MAST-Suche mit dem HPR- bzw. OPR-Motiv nach Proteinen gesucht. Hierbei wurden neben ausgewählten Vertretern der Alveolata und anderen Eukaryoten auch Bakterienvertreter untersucht. Gerade in den Spezies der Alveolata wurden sehr viele HPR-Proteine identifiziert. Die gefundenen HPR-Proteine zeigen wenig Überlappung mit OPR-Proteinen. Eine Ausnahme stellt *C. reinhardtii* dar. In diesem Organismus konnten mehr OPR- als HPR-Proteine identifiziert werden. Die hier gezeigte Darstellung dient der Zuordnung der Spezies und ist keine phylogenetische Abbildung der evolutionären Distanz. (Hillebrand et al. 2018)

OPR-Proteine identifiziert werden. Während in den Chromerida und den Dinoflagellaten eine große Zahl von HPR-Proteinen identifiziert wurde konnten in einem Vertreter der Ciliophora (*Tetrahymena thermophila*) auch nur 5 HPR-Proteine und keine OPR-Proteine identifiziert werden. Es ist spannend, dass innerhalb der Alveolata die Anzahl der HPR-Proteine scheinbar mit der Fragmentierung der rRNA Gene des Mitochondriums korreliert. Ob dies Zufall oder ein Hinweis auf die

potentielle Funktion der HPR-Proteine ist eine der interessantesten offenen Fragen dieser neuen Proteinfamilie.

### **3.4.2 HPR-Proteine können in vielen Taxa identifiziert werden, im Menschen werden FASTK-Proteine als HPR-Protein identifiziert**

Um die Verbreitung der HPR-Proteine außerhalb der Alveolata zu analysieren, wurden verschiedene prokaryotische und eukaryotische Genome durchsucht. Die Suche in Bakteriengenomen (*E. coli*, *Salmonella enterica* und *Bacillus subtilis*) blieb ebenso erfolglos wie die in Vertretern der Archaea (*Methanococcus maripaludis*, *Pyrococcus horikoshii*). In diesen wurden weder OPR- noch HPR-Proteine identifiziert. In Organismen der photoautotrophen *Stramenopiles* und *Cryptophyta* konnten HPR-Proteine identifiziert werden, jedoch in weit geringerer Zahl. In *Plasmodiophora brassicae*, einem einzelligen Pflanzenparasiten, der den meist heterotrophen *Cercozoa* zugeordnet wird, konnten 22 HPR-Proteine gefunden werden. Im Schwesterphylum der Alveolata, den Rhizaria (Vertreter *Reticulomyxa filosa*), konnten keine HPR- oder OPR-Proteine identifiziert werden. Im Phylum der *Archaeplastida*, in dem neben den höheren Pflanzen auch Algen zusammengefasst werden, können in einigen Spezies HPR-Proteine identifiziert werden. In Landpflanzen (vertreten durch *A. thaliana*) wurden keine HPR-Proteine gefunden. Der bereits mehrmals erwähnte Vertreter der Chlorophyta, *C. reinhardtii* ist die einzige untersuchte Spezies in der mehr OPR- als HPR-Proteine detektiert wurden. Die Anzahl der Proteine, die mit beiden Motiven identifiziert werden können, ist hier größer als in anderen Organismen.

Besonders interessant ist die Analyse des menschlichen Proteoms. Mit dem HPR-Motiv wurden 6 Proteine identifiziert, die zwischen 2 und 7 HPR-Motive zeigen. Von diesen 6 Proteinen weisen zwei (FASTKD3 und FASTKD1) allerdings einen e-Wert auf, der unterhalb der definierten Grenzen liegen. Diese Proteine gehören alle zur Familie der FASTK-Proteine (***Fas-activated-Serine/Threonine kinase-like***), einer kernkodierten Proteinfamilie, deren 6 Mitglieder mitochondriale Lokalisation zeigen (zusammengefasst in Jourdain et al. 2017). Eine schematische Darstellung der HPR-Motive innerhalb der FASTK-Proteine ist in Abbildung 14 gezeigt.

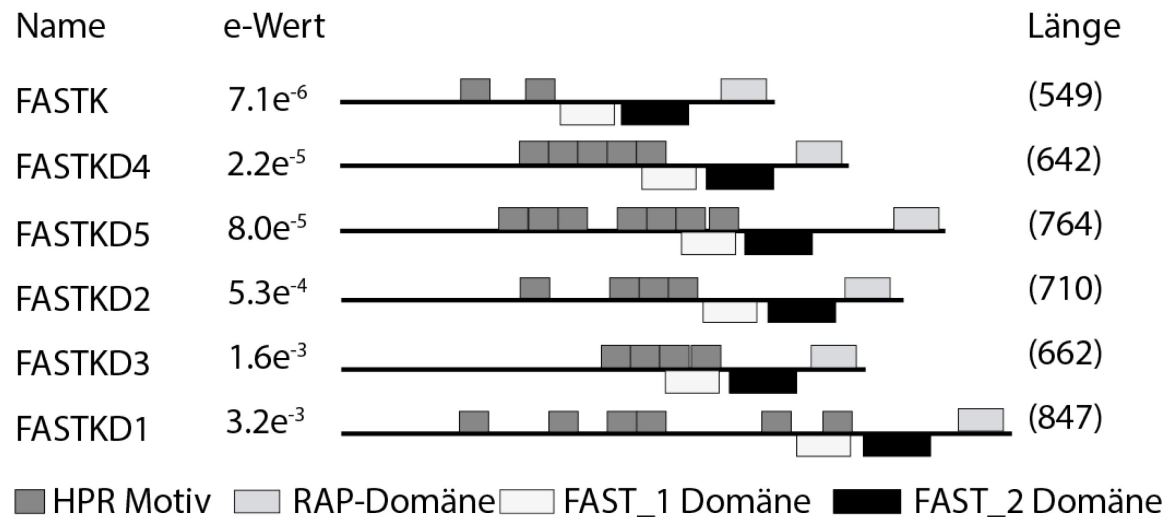


Abbildung 14: Schematische Übersicht der humanen Proteine mit HPR-Motiven. Die Suche mit dem HPR-Motiv (dunkelgraue Blöcke) identifizierte insgesamt 6 Proteine, die alle zur FASTK-Proteinfamilie gehören. Alle Mitglieder dieser Proteinfamilie besitzen neben eine C-terminale RAP-Domäne (hellgraue Blöcke), die zwei namensgebenden FAST\_Domänen (weiße bzw. schwarze Blöcke). Während die RAP- und FAST\_2-Domänen keine Überlappung mit den HPR-Motiven haben, überlagern sich die HPR-Motive und die FAST\_1 Domäne. Im Vergleich zu den in den Alveolata identifizierten HPR-Proteinen zeigen die FASTK-Proteine höhere e-Werte, was auf einen erhöhten Grad der Degeneration der HPR-Motivsequenz hindeuten kann. (Hillebrand et al. 2018)

Interessanterweise besitzen alle sechs FASTK-Proteine eine C-terminale RAP-Domäne und sind am RNA Metabolismus der humanen Mitochondrien beteiligt (Boehm et al. 2017, Antonicka und Shoubridge 2015). Die Suche mit dem OPR-Motiv im humanen Proteom identifizierte keine Proteine.

Diese Suchergebnisse zeigen eine deutlich abweichende Verteilung von HPR- und OPR-Proteinen. Die HPR-Proteine scheinen in den untersuchten Spezies wesentlich dominanter zu sein als die OPR-Proteine. Für die genaue Abgrenzung dieser zwei Proteinfamilien ist allerdings die Analyse von weiteren Spezies aus weiteren Phyla nötig. Eine Aussage über die Entstehung der HPR-Proteine erfordert ebenfalls weitergehende phylogenetische Analysen.

### 3.5 Funktionelle Untersuchung der neu entdeckten HPR-Proteine

Um eine funktionale Charakterisierung von ausgewählten HPR-Proteinen durchzuführen, wurden verschiedene genetische und biochemische Methoden verwendet. Die Proteine PBANKA\_051680 und PBANKA\_051380 aus *P. berghei* wurden für die folgenden Experimente besonders in den Fokus gestellt. Sie wurden in PbHPR1 (PBANKA\_051680) bzw. PbHPR2 (PBANKA\_051380) umbenannt. Diese beiden Proteine sind im Vergleich zu den anderen HPR-Proteinen kürzer (PbHPR1 ist 498 Aminosäuren lang; PbHPR2 ist 441 Aminosäuren lang, siehe Abbildung 12). Die Proteingröße der beiden HPR-Proteine spielt vor allem bei der rekombinanten Expression in *E. coli* eine Rolle. Die Expression von stark helikalen Proteinen gestaltet sich in *E. coli* schwierig, da die rekombinanten Proteine zur Bildung von Aggregaten neigen und dadurch schwerer löslich sind. Eine Verbesserung der Expression kann durch die Reduktion der Proteingröße erreicht werden (Prof. A. Barkan, *University of Oregon*, persönliche Kommunikation). Beide Proteine besitzen eine vergleichbare Anzahl von HPR-Motive in direkter Folge (siehe Abbildung 12). Obwohl die Proteine in Bezug auf ihre Größe und die HPR-Motive vergleichbar sind, enthält PbHPR2 eine C-terminale RAP-Domäne. In PbHPR1 wurden keine weiteren Domänen identifiziert.

#### 3.5.1 Die HPR-Proteine sind dominiert durch helikale Elemente, die auch außerhalb der HPR-Motive liegen

Wie in der Einleitung erwähnt, zeichnen sich OPR- und PPR-Proteine durch eine von  $\alpha$ -Helices dominierte Sekundärstruktur aus. Das gesamte PPR-Protein formt eine solenoide Struktur. Um einen tieferen Einblick in die potentielle Sekundärstruktur der HPR-Proteine zu erhalten wurden PbHPR1 und PbHPR2 mit I-TASSER<sup>9</sup> (*Iterative Threading ASSEmbly Refinement*, Roy et al. 2010) modelliert. Hierzu wurde die vollständige Proteinsequenz inklusive des N-terminalen Lokalisationspeptides genutzt (PlasmoDB). Die generierten Sekundärstrukturen sind in Abbildung 15 dargestellt. Beide Proteine bestehen fast ausschließlich aus  $\alpha$ -Helices. Diese helikalen Elemente sind auch außerhalb der identifizierten HPR-

---

<sup>9</sup> <https://zhanglab.ccmb.med.umich.edu/I-TASSER/>



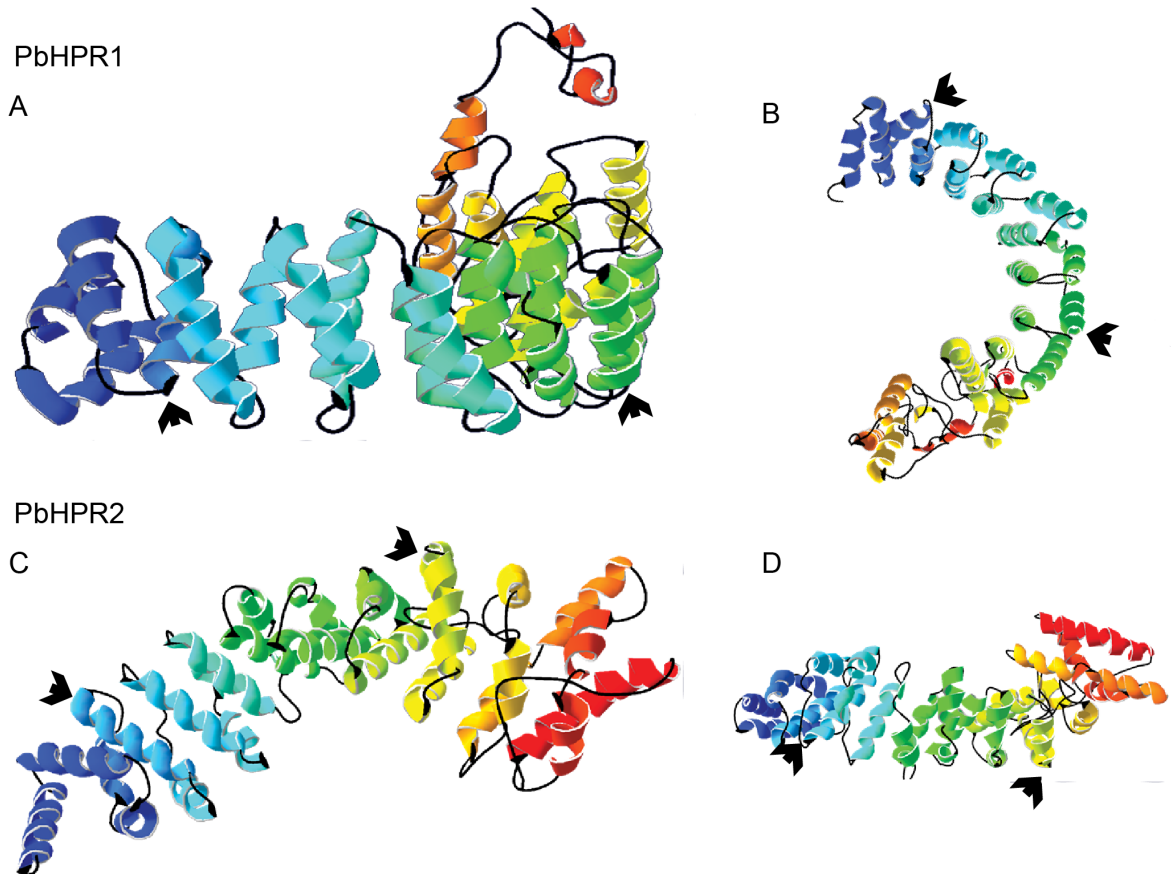


Abbildung 15: Die Strukturvorhersage von PbHPR1 und PbHPR2. Die durch I-TASSER erstellten Strukturmodelle von beiden Proteinen sind von  $\alpha$ -Helices dominiert. Der N-Terminus ist in blau eingefärbt, der C-Terminus in Rot. Teilweise befinden sich die helikalen Elemente außerhalb der identifizierten HPR-Motive. Anfang und Ende der HPR-Motivakkumulation sind mit schwarzen Pfeilen markiert. In PbHPR1 wurde für die Darstellung **(A)** ein kleiner unstrukturierter Bereich am C-Terminus entfernt. In **(A)** und **(B)** ist PbHPR1 aus zwei verschiedenen Winkeln gezeigt, die halbkreisförmige Gesamtstruktur ist deutlich sichtbar **(B)**. PbHPR2 formt eine lineare Gesamtstruktur **(C und D)**, wobei das Protein eine leichte Rotation um die Längsachse des Proteins zeigt. (Hillebrand et al. 2018)

Motive zu finden (siehe auch Anhang Abbildung 2). Es ist zu vermuten, dass diese helikalen Elemente ebenfalls HPR-Motive bilden, jedoch mit dem bisherigen Konsensusmotiv nicht identifiziert werden (siehe Abbildung 15, schwarze Pfeile). Das PbHPR1 Protein formt eine halbkreisartige Gesamtstruktur und ähnelt damit der Struktur von PPR-Proteinen (Abbildung 15 B). Während in PPR-Proteinen die Anordnung entlang der Proteinachse sehr regelmäßig ist, scheint in PbHPR1 die Anordnung der helikalen Elemente etwas weniger regelmäßig. Das Modell von PbHPR2 deutet auf eine eher lineare als solenoide Gesamtstruktur hin (Abbildung 15 C und D). Entlang der Proteinachse ist eine leichte Rotation der Gesamtstruktur

zu erkennen. Aufgrund von fehlenden OPR-Strukturdaten ist es nicht möglich die näher verwandten OPR-Proteine mit den HPR-Strukturen zu vergleichen.

### 3.5.2 Ausgewählte HPR-Proteine lokalisieren *in vivo* in Mitochondrien von *P. berghei*

Die vorhergesagte Lokalisation der HPR-Proteine deutete auf eine mitochondriale Lokalisation hin. Da die mitochondrialen Transitpeptide in *Plasmodium* sehr variabel sind (siehe bspw. Bender et al. 2003) und so in den genutzten Vorhersagealgorithmen die Möglichkeit für falsch-positiven Ergebnissen erhöht ist, wurden 7 HPR-Proteine in *P. berghei* genetisch markiert und durch Fluoreszenzmikroskopie analysiert. Die Proteinmarkierung erfolgte durch homologe Rekombination am C-Terminus des HPR-Proteins mit dem rot-fluoreszierenden Protein *mCherry*. Der genutzte Transfektionsvektor beinhaltet zusätzlich die kodierende Sequenz für ein mitochondrial lokalisierendes GFP (**g**reen **f**luorescent

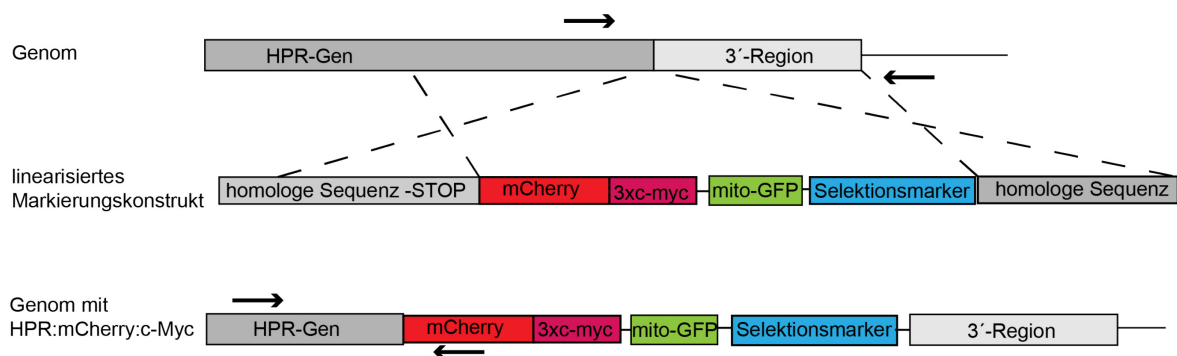


Abbildung 16: Darstellung der verwendeten Markierungskonstrukte. Die Fluoreszenzmarkierung ausgewählter Proteine erfolgte durch doppelte homologe Rekombination. Das Markierungskonstrukt enthält die kodierende Sequenz für ein Fluorophor (mCherry) direkt gefolgt von der Sequenz einer biochemischen Markierung (c-Myc). Die Selektion in den Parasiten erfolgte über die Resistenzkassette (Selektionsmarker), die eine Resistenz gegen Pyrimethamin vermittelt. Um eine Verfolgung des Mitochondriums zu ermöglichen wurde ein mitochondrial lokalisiertes GFP verwendet (mitoGFP). Um die homologe Rekombination zu ermöglichen, wurde der 3'-Bereich des HPR-Gens durch PCR amplifiziert und gleichzeitig das Stopcodon entfernt und zusammen mit der ebenfalls amplifizierten 3'-nichtkodierenden Region verwendet. Mit diesem linearisierten Konstrukt wurden Wildtyp-Parasiten von *P. berghei* ANKA transfiziert. Der Integrationserfolg wurde durch PCR verfolgt, die verwendeten Primer sind als schwarze Pfeile eingezeichnet.

protein), um in der Mikroskopie das Mitochondrium ebenfalls fluoreszent zu markieren. Die Transfektionsvektoren wurden, wie in 2.2.34 beschrieben, erstellt. Ein schematischer Aufbau des Vektors ist in Abbildung 16 gezeigt. Die Transfektion der Wildtyp-Parasiten von *P. berghei* ANKA erfolgte wie in 2.2.33 beschrieben. Der Integrationserfolg wurde durch PCR-Amplifikation mit Kombinationen aus vektorspezifischen und genomspezifischen Oligonukleotiden analysiert. Die Analyse der PCR auf einem 1%igen Agarosegel dokumentiert bei allen gezeigten Konstrukten eine korrekte Integration ins Genom (siehe Anhang Abbildung 1). Die korrekte Integration der Konstrukte für die HPR-Proteine PbHPR1 und PBANKA\_081630 wurden in der Masterarbeit von Martin Almendinger unter meiner Betreuung durchgeführt. Die korrekte Integration dieser Konstrukte wurde ebenfalls bestätigt.

Nach der Transfektion wurden die Parasiten in NMRI-Mäuse injiziert. Nach 5 Tagen wurde das infizierte Mäuseblut mikroskopiert. Die Mikroskopie wurde mit lebenden Zellen ohne Fixierung durchgeführt. Um die Zellkerne zu identifizieren, wurde die DNA mit dem Farbstoff Hoechst 33342 (Thermo Fisher) angefärbt. Die Fluoreszenzmarkierung wurde für 7 Proteine (PBANKA\_040990 Ortholog zu PF3D7\_0311900, PBANKA\_143670 Ortholog zu PF3D7\_1221500, PBANKA\_111800 Ortholog zu PF3D7\_0618800, PBANKA\_142250 Ortholog zu PF3D7\_0815100, PBANKA\_081630 Ortholog zu PF3D7\_0915300, PBANKA\_051380 Ortholog zu PF3D7\_1029800, PBANKA\_051680 Ortholog zu PF3D7\_1033000) durchgeführt. Von 5 dieser Proteine sind repräsentative mikroskopische Bilder in Abbildung 17 gezeigt. Das rote Signal für das fluoreszenzmarkierte Protein ist bei allen untersuchten Proteinen in einer distinkten Struktur innerhalb der Zelle zu erkennen, während die DNA des Zellkerns separat vorliegt. Die Form des roten Fluoreszenzsignals überlappt vollständig mit dem grünen Signal des mitochondrialen GFPs. Um ein zum GFP-Signal vergleichbar starkes mCherry-Signal zu erhalten, musste jedoch die Expositionszeit verlängert werden, sodass ein relativ starkes Hintergrundsignal erkennbar ist. Dies deutet darauf hin, dass die HPR-Proteine, ähnlich wie die meisten pflanzlichen PPR-Proteine, ein relativ niedriges Expressionslevel in Blutstadien haben. Die Lokalisation im Mitochondrium konnte bei allen untersuchten Proteinen bestätigt werden. In keinem Fall wurde ein Überlapp zwischen rotem Signal und Kern

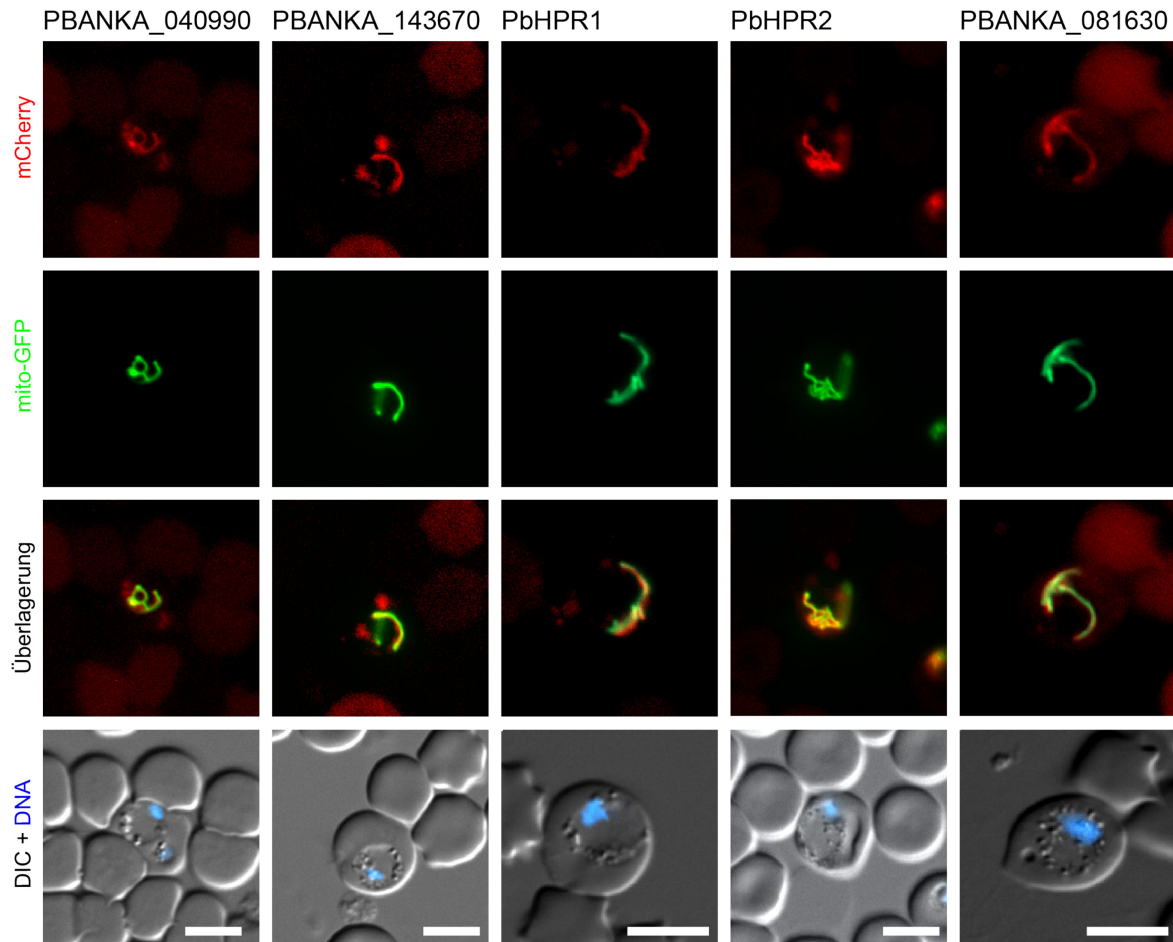


Abbildung 17: Mikroskopische Analyse ausgewählter, fluoreszenzmarkierter HPR-Proteine. Ausgewählte HPR-Proteine wurden mit fluoreszenten Proteinen markiert und mikroskopisch analysiert. Das Signal der Proteine ist in Rot (mCherry) dargestellt, während das markierte Mitochondrium in Grün gezeigt ist (mito-GFP). Die Färbung der DNA erfolgte durch Hoechst 33342 (blau). Es ist eine deutliche Überlagerung der roten und grünen Signale zu erkennen, während die DNA in einem distinkten Kompartiment angefärbt wird (Zellkern). Da für die Aufnahmen der roten Fluoreszenz eine längere Exposition verwendet werden musste, ist ein deutliches Hintergrundsignal zu erkennen. Die mitochondriale Lokalisation der HPR-Proteine konnte somit bestätigt werden. Längenmarker = 5  $\mu$ m. (teilweise Hillebrand et al. 2018)

detektiert. Bei PBANKA\_142250 ist neben dem mitochondrialen Signal ein weiteres Signal in zwei Punkten außerhalb von Kern und Mitochondrium zu erkennen. Ob es sich dabei um ein Artefakt handelt oder ob das Protein partiell auch außerhalb des Mitochondriums vorliegt ist unklar.

Das HPR-Protein PBANKA\_020810 (Ortholog zu PF3D7\_0105200) wurde in der Gruppe von Dr. Ross Waller (University of Cambridge, England) untersucht und die mitochondriale Lokalisation dokumentiert (persönliche Kommunikation). Diese Ergebnisse können also die Lokalisationsvorhersage bestätigen, sodass der Großteil der HPR-Proteine als mitochondriale Proteine angesehen werden kann. Für die Proteine, die potentiell im Apicoplasten vorliegen, steht eine Bestätigung der Lokalisation noch aus.

### **3.6 Deletion von HPR-Proteinen deuten auf essentielle Funktion der HPR-Proteine in den Blutstadien hin**

Zur weiteren Untersuchung der HPR-Proteine in *P. berghei* wurde für 15 Proteine der Versuch unternommen, Deletionsmutanten zu erstellen. Die Deletion der HPR-Gene erfolgte durch doppelte homologe Rekombination. Der Deletionsvektor wurde wie in 2.2.35 beschrieben erstellt und enthält homologe Sequenzen zu den Sequenzen unmittelbar stromauf und stromab der HPR-kodierenden Sequenz (siehe Schema in Abbildung 18 A). Die Transfektion des linearisierten Konstruktes erfolgte wie in 2.2.33 beschrieben. Das Blut, der mit den Parasiten infizierten Mäuse, wurde nach 5-8 Tagen abgenommen und die parasitäre genomische DNA isoliert (siehe 2.2.22). Der Erfolg der Deletion wurde durch PCR überprüft. Es wurden Oligonukleotide spezifisch für den Deletionsvektor in Kombination mit genspezifischen Oligonukleotiden genutzt, die außerhalb der Rekombinationssequenzen lagen (siehe Abbildung 18 A, schwarze Pfeile).

Für 13 der 15 HPR-Proteine konnten keine Deletionsmutanten generiert werden. Auch die Wiederholung der Deletionsversuche brachte keinen Erfolg. In der PCR-Analyse der transfizierten Parasiten konnte ausschließlich das Wildtyp-Allel detektiert werden (eine Auswahl der PCR-Ergebnisse ist in Abbildung 18 B gezeigt; eine Liste der mutierten Gene findet sich in der nachfolgenden Tabelle 6). In der Tabelle sind ebenfalls die ermittelten Phänotypen in genomweiten Deletionsstudien für *P. berghei* (Bushell et al. 2017) und *P. falciparum* (Zhang et al. 2018) aufgeführt. In diesen genomweiten Studien werden HPR-Proteine ebenfalls meist als essentiell bestätigt.

## Ergebnisse

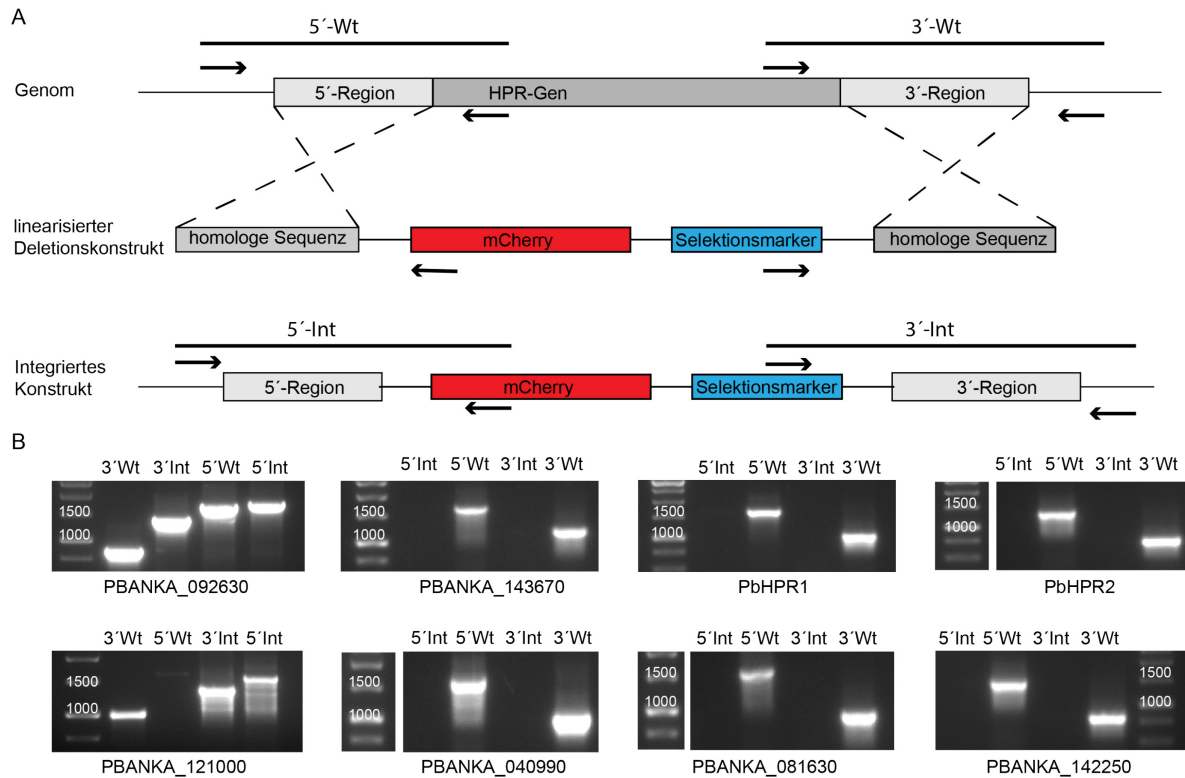


Abbildung 18: Erstellung von Deletionsmutanten für HPR-Proteine. **A)** Für die Deletion von ausgewählten HPR-Proteinen wurden Deletionskonstrukte generiert, die neben einem cytosolisch lokalisierten Fluoreszenzmarker (mCherry) und einem Selektionsmarker (DHFR), der die Resistenz gegen Pyrimethamin vermittelt, genspezifische, homologe Sequenzen enthalten. Das Deletionskonstrukt wird durch Endonukleasen linearisiert und für die Transfektion von *P. berghei* ANKA Wildtyp-Parasiten verwendet. Im Falle der erfolgreichen homologen Rekombination wird das HPR-Gen aus dem Genom entfernt und stattdessen das Deletionskonstrukt ins Genom integriert. Die erfolgte Integration des Konstruktes wurde durch PCR-Methoden überprüft. Die verwendeten Primer (schwarze Pfeile) und die erwarteten Amplikons (schwarze Balken) sind ebenfalls gezeigt. **B)** In der Überprüfung konnte für zwei HPR-Gene eine Deletion detektiert werden (PBANKA\_092630 und PBANKA\_121000). Sowohl für die unveränderte Genomsituation (Spur 3'- bzw. 5'-Wt) als auch für die Integration des Konstruktes (3'- bzw. 5'-Int) waren Amplifikate zu erkennen. Dies deutet auf eine Mischpopulation zwischen Deletionsmutante und unverändertem Wildtyp hin. Die anderen 6 gezeigten Konstrukte zeigen nur das unveränderte Wildtyp-Genom an. Hier sind 6 Beispiele für negative Integration gezeigt. Insgesamt wurden 13 vermutlich essentielle HPR-Gene identifiziert (siehe Tabelle 6).

In zwei Deletionsversuchen konnte die Integration des Deletionskonstruktes nachgewiesen werden. Die Proteine PBANKA\_121000 und PBANKA\_092630 scheinen also für den Parasiten nicht essentiell zu sein (Abbildung 18 B). Wie aus

der Genotypisierung ersichtlich, beinhaltet die Parasitenpopulation sowohl Wildtyp als auch Deletionsmutanten. Um eine reine Parasitenzelllinie zu erhalten, wurden die Parasiten durch FACS-Methoden (*Fluorescence-activated cell sorting*) anhand des Fluoreszenzsignals der mCherry-Proteine sortiert und anschließend erneut genotypisiert. Für die Deletion des PBANKA\_121000 Proteins konnte durch FACS eine reine Deletionszelllinie generiert werden. In der Deletion des HPR-Proteins PBANKA\_092630 wurde auch nach zwei Sortierdurchläufen immer noch eine Mischpopulation aus Wildtyp und Deletionsmutante festgestellt. Aus Zeitgründen konnte in dieser Arbeit die Analyse dieser Mutanten nicht mehr durchgeführt werden. Die Charakterisierung der Deletionsmutanten und besonders die Analyse der mitochondrialen Transkripte dieser Mutanten ist eine spannende Möglichkeit den HPR-Proteinen eine Funktion zuzuordnen.

Tabelle 6: Liste der Deletionsversuche von HPR-Proteinen. Auch aufgeführt sind die Phänotypen der HPR-Proteine in publizierten, genomweiten Deletionsstudien.

HPR-Protein	erfolgreiche Deletion	Phänotyp in Bushell et al. 2017	Ortholog in <i>P. falciparum</i>	Phänotyp in Zang et al. 2018
PBANKA_020810	-	essentiell	PF3D7_0105200	Deletion möglich
PBANKA_080730	-	Deletion möglich	PF3D7_0317900	essentiell
PBANKA_143670	-	essentiell	PF3D7_1221500	essentiell
PBANKA_051680	-	essentiell	PF3D7_1033000	essentiell
PBANKA_142250	-	essentiell	PF3D7_0815100	essentiell
PBANKA_092630	+	nicht enthalten	PF3D7_1122000	essentiell
PBANKA_124410	-	ungenügende Daten	PF3D7_0529700	essentiell
PBANKA_081630	-	nicht enthalten	PF3D7_0915300	essentiell
PBANKA_144130	-	nicht enthalten	PF3D7_1226500	essentiell
PBANKA_040990	-	nicht enthalten	PF3D7_0311900	essentiell
PBANKA_051380	-	nicht enthalten	PF3D7_1029800	essentiell
PBANKA_082510	-	nicht enthalten	PF3D7_0924200	essentiell
PBANKA_112060	-	nicht enthalten	PF3D7_0621700	essentiell
PBANKA_121000	+	Deletion möglich	PF3D7_1011600	Deletion möglich
PBANKA_145840	-	nicht enthalten	PF3D7_1245200	essentiell

### 3.7 Analyse der RNA-Bindung von PbHPR1 und PbHPR2 *in vitro*

PPR- und OPR-Proteine binden direkt RNA. Zur Charakterisierung der biochemischen Eigenschaften der HPR-Proteine, sollte in *in vitro* Analysen mit

rekombinant hergestellten HPR-Proteinen ihre potentielle Fähigkeit RNA zu binden untersucht werden. Da diese Analyse die genetischen Experimente flankieren sollten, wurden sie mit HPR-Proteinen von *P. berghei* durchgeführt.

### 3.7.1 Expression von rekombinantem PbHPR1 und PbHPR2 in *E. coli*

PbHPR1 und PbHPR2 sollten als Fusionsproteine mit einem N-terminalen MBP (Maltose Bindeprotein) durch heterologe Expression in *E. coli* hergestellt werden. Hierzu wurde das pMAL™ *Protein Fusion and Purification System* (NEB) verwendet, welches in der Vergangenheit von anderen Arbeitsgruppen erfolgreich für die Expression von PPR-Proteinen verwendet worden war. Helikale Proteine neigen zur Aggregation. Das verwendete MBP-Protein erhöht jedoch die Löslichkeit. Um das mitochondriale Lokalisationssignal der HPR-Proteine zu entfernen, wurden bei der Amplifikation der HPR-Gene im N-Terminus des Proteins durch die Wahl der Oligonukleotide einige Basen entfernt. Da die Vorhersage der mitochondrialen Translokationssequenz für *P. berghei* Proteine mit schwierig ist (Bender et al. 2003), wurden die orthologen Proteine für PbHPR1 und PbHPR2 aus verschiedenen *Plasmodium* Spezies genutzt, um basierend auf der Sequenzkonservierung in Kombination mit der Position der HPR-Motive die Translokationssequenz zu identifizieren (siehe Anhang Abbildung 3). Die Konservierung wurde mit der Software *CLC-Workbench* (Qiagen) bestimmt. Basierend auf diesen Daten wurden Oligonukleotide erstellt, die kodierenden Sequenzen der beiden Proteine durch PCR amplifiziert und gleichzeitig die Erkennungssequenzen für die Restriktionsendonukleasen *BamHI* und *Sall* eingefügt. Da beide Gene keine Introns enthalten wurden die PCR mit genomischer DNA als Vorlage durchgeführt. Nach dem Verdau mit diesen Enzymen wurden die Amplifikate mit dem Expressionsvektor pMAL-TEV ligiert. Die Expressionsvektoren wurden mit Sanger-Sequenzierung überprüft. Zur Expression der rekombinanten Proteine wurden chemisch-kompetente Zellen (*E. coli* RIPL) mit den Expressionskonstrukten transformiert. Dieser Stamm zeichnet sich durch zusätzliche Plasmide aus, die für tRNAs mit selteneren Codons kodieren.

Nach einem Pilotexperiment zeigte sich, dass die Expression von MBP:PbHPR2 mit der nativen Sequenz erfolglos war. Da das Genom von *P. berghei* sehr AT-reich ist, wurde die kodierende Sequenz des Proteins für die Expression in



## Ergebnisse

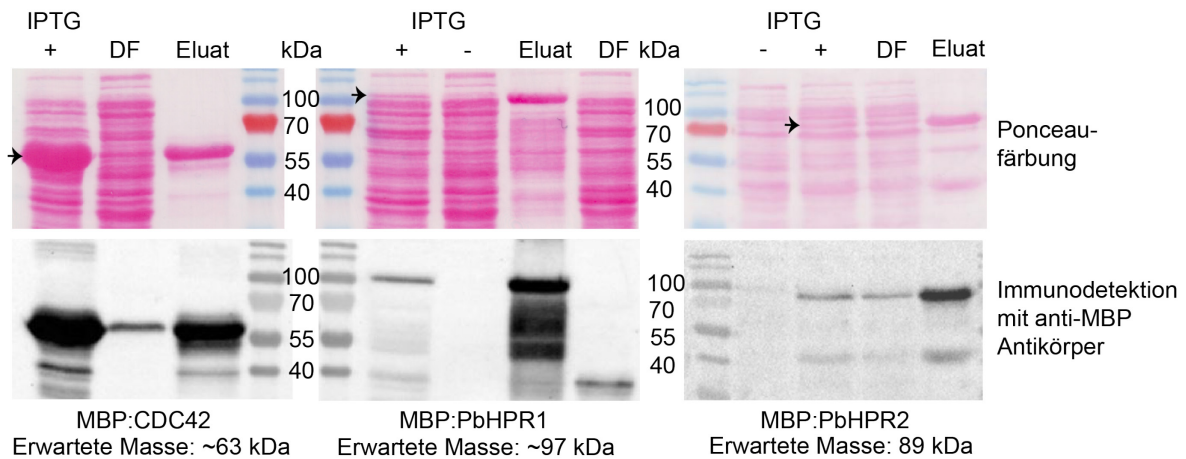


Abbildung 19: Heterologe Expression von PbHPR1 und PbHPR2. In der oberen Reihe ist die unspezifische Proteinfärbung der Membranen mit Ponceau gezeigt, in der unteren Reihe die Immundetektion im *Western-blot* mit einem Anti-MBP-Antikörper. Die Expression erfolgte als Fusionsproteine mit N-terminalem MBP in *E. coli* RIPL Zellen. Die Induktion der Expression erfolgte durch IPTG. Während keine Expression der Fusionsproteine ohne IPTG stattfindet (Spur IPTG -, nur bei MBP:HPR Fusionsproteinen), ist bei der Induktion mit IPTG im Zellextrakt ein Signal in der berechneten Größe von MBP:PbHPR1 (Mitte) und MBP:PbHPR2 (Rechts) zu sehen. Diese Banden sind in der unspezifischen Ponceaufärbung erkennbar (schwarze Pfeile). Die folgende chromatographische Reinigung des Fusionsproteins erfolgt durch die Bindung des MBP an Amylosesäulen. Der Durchfluss dieser Chromatographie (DF) zeigt nur schwache Signale in der Immundetektion der Fusionsproteine, während die meisten zellulären Proteine im Durchfluss enthalten bleiben. In den Elutionsfraktionen (Eluat) ist eine deutliche Anreicherung des Proteins zu erkennen. Die meisten unspezifischen Proteine wurden gewaschen. In der Immundetektion sieht man vor allem in der Elution von MBP:PbHPR1 Signale zwischen dem freien MBP und dem Fusionsprotein, die wahrscheinlich C-terminalen Degradationsprodukten von MBP:PbHPR1 entsprechen. (Hillebrand et al. 2018)

*E. coli* optimiert und anschließend kommerziell synthetisiert. Für die Optimierung wurde die Software GeneArt (Thermo Fisher) verwendet. Die Synthese erfolgte kommerziell über Genewiz inc. (GENEWIZ European Headquarter & Leipzig Laboratory). Bei der Synthese wurden an den entsprechenden Enden die Erkennungssequenzen für die Restriktionsendonukleasen eingefügt (*Bam*HI am 5'-Ende und *Sal*I am 3'-Ende). Die synthetisierten DNA-Moleküle wurden nach Herstellerangaben in deionisiertem Wasser zu einer Konzentration von 100 µg/µl gelöst. Anschließend erfolgten der Restriktionsverdau und die Ligation in den

Expressionsvektor. Auch diese Vektoren wurden mittels Sanger-Sequenzierung überprüft.

Die Expression und anschließende Isolation der Proteine erfolgte wie in 2.2.29 beschrieben. Der Expressions- und Anreicherungs-erfolg wurde durch Immundetektion des MBP in einem *Western-Blot* (*Anti-MBP Monoclonal Antibody*, NEB) überprüft (siehe Abbildung 19). Die Proteinkonzentrationen wurden mittels Bradford-Bestimmung (siehe 2.2.26) ermittelt und das Proteingel mit 25 µg vom jeweiligen Zelllysate und dem Durchfluss der Chromatographie beladen. Vom eluierten Protein wurden jeweils 5 % des Elutionsvolumens (200 µl) aufgetragen. Als Kontrolle wurde gleichzeitig ein Fusionsprotein aus MBP und CDC42 (*cell division control protein 42*) aus *Drosophila melanogaster* exprimiert. Die Immundetektion der Zelllysate dokumentiert, dass bei allen drei Proteinen durch die Zugabe von IPTG ins Medium die Expression der Fusionsproteine aktiviert wird (Abbildung 19, Spuren +IPTG). Sowohl im Gesamtzellextrakt als auch in der aufgereinigten Proteinfraction sind zusätzliche Banden in der Größe von freiem MBP (42 kDa Proteinmasse) zu erkennen. Das Kontrollprotein und MBP:PbHPR2 zeigen nach der chromatographischen Reinigung kaum Degradationsprodukte. In PbHPR1 sind in der Immundetektion zusätzliche Signale zwischen dem intakten Fusionsprotein und MBP zu sehen. Diese repräsentieren vermutlich Degradationsprodukte des Fusionsproteins (Abbildung 19 mittleres Bild, Spur Eluat). Nach der Durchführung der chromatographischen Reinigung des Fusionsproteins wurden die Elutionsfraction ebenfalls durch *Western-blot* untersucht. Die Elution der Proteine von der Amylose erfolgte innerhalb der ersten 3 Fractionen (jeweils 1 ml). Diese Fractionen wurden vereinigt und durch die Zentrifugation durch *Amicon® Ultra Centrifugal Filters* (Ausschlussgröße 50 kDa) auf 150µl eingeengt. Die gereinigten Proteine wurden zur Konservierung mit 1 Vol. 50 % Glycerol vermischt und entweder *in vitro* RNA-Bindestudien verwendet oder bei -20°C gelagert. Anhand der immunologischen Analyse konnten beide HPR-Proteine in *E. coli* exprimiert und durch chromatographische Methoden angereichert werden (siehe Abbildung 19). In der Elutionsfraction des PbHPR1-Proteins sind deutlich Degradationsprodukte erkennbar. In der Expression des Kontrollproteins (MBP:CDC42) und in der Expression des PbHPR2-Proteins wurden deutlich weniger Degradationen festgestellt.

### 3.7.2 Herstellung der mitochondrialen RNAs durch *in vitro* Transkription

Aufgrund der geringen Größe des mitochondrialen Genoms ist es möglich, durch *in vitro* Transkription das gesamte Transkriptom des Organells abzubilden. Das mitochondriale Genom von *P. berghei* wurde dafür in 4 überlappende Fragmente unterschiedlicher Größe aufgeteilt und durch PCR amplifiziert (siehe Abbildung 20). Um eine strangspezifische Transkription zu ermöglichen, wurde bei der PCR jeweils ein Strang mit einer Promotorsequenz der T7-RNA-Polymerase versehen. Die PCR Produkte wurden aufgereinigt und als Vorlage in einer *in vitro* Transkription mit T7-DNA abhängigen RNA-Polymerase (Thermo Fisher) verwendet. Die hergestellten Transkripte hatten eine Länge zwischen 1200 und 1977 nt (siehe Abbildung 20). Nach erfolgter Transkription wurde die DNA-Vorlage durch die Zugabe einer DNase (4 U TURBO™ DNase, Invitrogen) abgebaut und die RNA anschließend durch RNA-Extraktion isoliert. Aufgrund der unterschiedlichen Größe der Fragmente ergibt sich bei der Gelelektrophorese das in Abbildung 20 (rechtes Bild) gezeigte Bandenmuster. Die so generierte RNA wurde zusammen mit den rekombinanten HPR-Proteinen in RNA-Bindestudien verwendet.

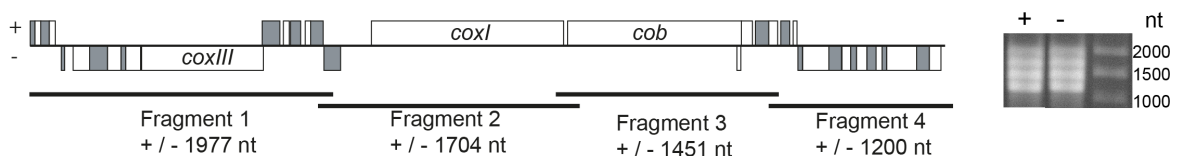


Abbildung 20: Aufteilung des mitochondrialen Genoms für die Herstellung von *in vitro* Transkripten, welche das gesamte Genom abdecken. Das mitochondriale Genom von *P. berghei* wurde in 4 unterschiedliche Fragmente aufgeteilt und durch PCR amplifiziert. Die Fragmente wurden während der PCR strangspezifisch mit einer T7-Promotorsequenz versehen und als Vorlage in der *in vitro* Transkription verwendet. Die generierten RNAs wurden isoliert und in einer denaturierenden RNA-Gelelektrophorese getrennt (rechtes Bild). Die so generierten Fragmente wurden für RNA-Bindestudien genutzt. (Hillebrand et al. 2018)

### 3.7.3 Rekombinantes PbHPR1-Protein bindet *in vitro* unspezifisch mitochondriale RNAs

Die gereinigten HPR-Proteine und das Kontrollprotein wurden in verschiedenen Puffern mit den *in vitro* Transkripten jeweils eines Stranges inkubiert, um die potentielle Interaktion zwischen RNA und HPR-Proteinen zu untersuchen.

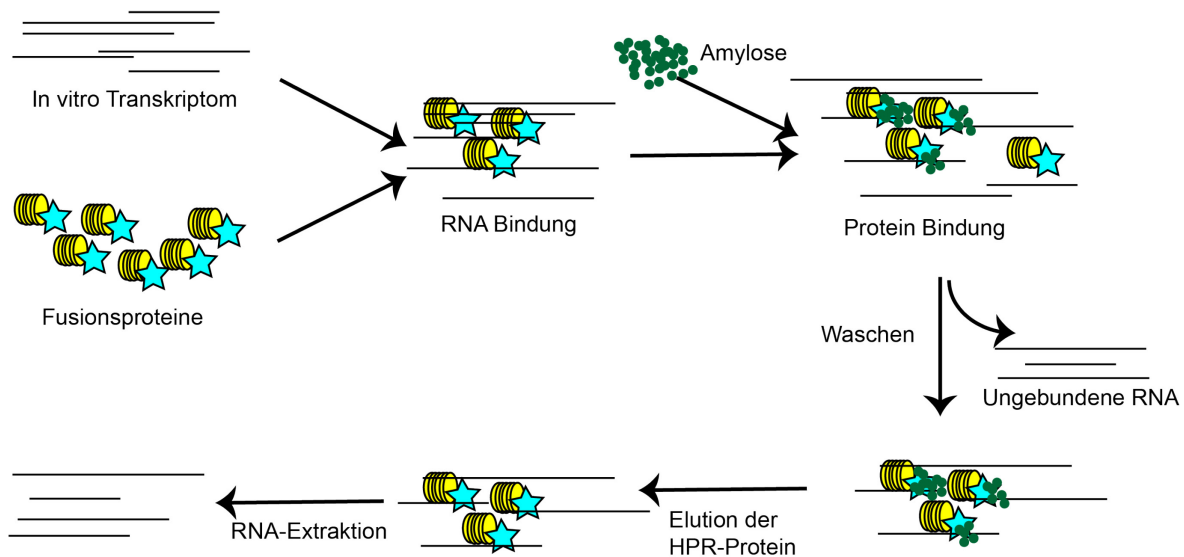


Abbildung 21: Schematischer Ablauf der RNA-Bindestudie mit rekombinanten MBP:HPR-Fusionsproteinen. Die rekombinanten Fusionsproteine (HPR-Protein = Gelb; MBP = Stern) wurden zusammen mit den durch *in vitro* Transkription hergestellten RNAs inkubiert. Anschließend wurden die Proteine, durch die Bindung des MBP an Amylose, isoliert. Unspezifische Interaktion der RNA mit der Amylose oder den Proteinen wurde durch Waschen minimiert. Die mit den Proteinen eluierte RNA wurde anschließend extrahiert und in RNA-Gelelektrophorese mit angeschlossener radioaktiver Detektion analysiert. Um die Doppelstrangbildung der RNA zu minimieren, wurde die RNA strangspezifisch eingesetzt. Die Fusionsproteine aus MBP:CDC42 wurden als Kontrolle gleichbehandelt.

Ein Schema des Experiments ist in Abbildung 21 gezeigt. Die Fusionsproteine (520 nmol) wurden mit der RNA (6 µg von jedem Fragment) für 20 Minuten bei Raumtemperatur inkubiert. Durch die Zugabe von Amylose wurden die Fusionsproteine gebunden und durch Zentrifugation der Überstand und das Amylosematerial getrennt. Der Erfolg der Präzipitation wurde durch immunologische Detektion verfolgt. Für diese wurden jeweils 10 % des Volumens der einzelnen Schritte in einem SDS-PAGE Gel aufgetrennt und in der anschließenden Detektion die Fusionsproteine durch MBP-spezifische Antikörper detektiert (siehe 2.2.28). Die immunologische Überprüfung der Präzipitation zeigt, dass geringe Mengen des Proteins präzipitiert werden konnten, jedoch verbleibt ein größerer Teil im Überstand. Dies gilt sowohl für die Kontrolle (MBP:CDC42), als auch für die rekombinanten HPR-Fusionsproteine (siehe Abbildung 22 A).

Der Überstand der Präzipitate wurde abgenommen und die RNA extrahiert. Unspezifisch an das Amylosematerial gebundene RNA sollte durch mehrere

Waschschritte minimiert werden. Anschließend wurden die Proteine durch maltosehaltigen Puffer (10 mM Endkonzentration Maltose) eluiert und die RNA aus dem Eluat extrahiert. Die eingesetzte RNA, die ungebundenen RNA des Überstand und die zusammen mit den Proteinen präzipitierte RNA wurden anschließend durch YelZol-Extraktion isoliert (siehe 2.2.12). Auf einem denaturierenden Agarose-Gel wurden die extrahierten RNAs elektrophoretisch getrennt und auf eine Membran transferiert. Zur Detektion der transferierten RNA wurden komplementäre

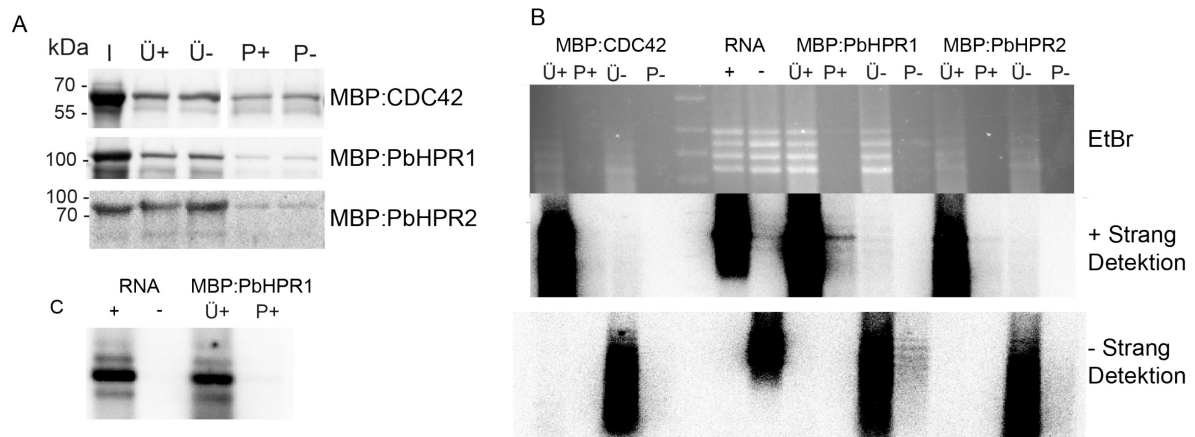


Abbildung 22: RNA-Bindestudien mit rekombinanten HPR-Proteinen. Die RNA Bindestudien wurden mit strangspezifischen Transkripten des mitochondrialen Genoms von *P. berghei* durchgeführt (als + und – bezeichnet, siehe Abbildung 20) **A**) Die Präzipitation der Fusionsproteine wurde durch Immundetektion mit Anti-MBP Antikörpern durchgeführt. Die eingesetzten Proteine (I) wurden zusammen mit den strangspezifischen Transkripten inkubiert und anschließend durch Amylose isoliert. Nur ein Teil der Proteine konnte isoliert werden (P+ bzw. P-) während ein größerer Teil im Überstand verbleibt (Ü+ bzw. Ü-). **B**) Die isolierte RNA wurde auf einem 1 % denaturierenden Agarosegel aufgetrennt und mit radioaktiv markierten, strangspezifischen Oligonukleotide detektiert. Die eingesetzten Transkripte (RNA + bzw. -) und die Überstände (Ü+ bzw. Ü-) sind deutlich überexponiert, um die geringen RNA-Mengen in der präzipitierten Proteinfraction darzustellen (P+ bzw. P-). Während sowohl im Plusstrang als auch im Minusstrang von PbHPR1 RNA detektiert wird, sind weder in den Präzipitationen der Kontrolle (MBP:CDC42), noch des zweiten HPRs (MBP:PbHPR2) RNA Signale zu erkennen. In der Präzipitation von PbHPR1 scheint eine der Banden des Plusstranges dominanter zu sein. Jedoch ist in **C**) die gleiche Detektion mit geringerer Exposition gezeigt. Es zeigt sich, dass sowohl in den eingesetzten Transkripten als auch im Überstand diese Bande dominanter ist. Dies deutet auf technische Gründe bei der radioaktiven Markierung der Oligonukleotide hin. (Hillebrand et al. 2018)

Oligonukleotide für die vier RNA-Fragmente des jeweiligen Genomstranges radioaktiv markiert und zur Detektion verwendet.

Bei der radioaktiven Detektion der strangspezifischen Transkripte konnte für das Kontrollprotein MBP:CDC42 keine mit den Proteinen co-präzipitierte RNA detektiert werden (Abbildung 22 B, P+ und P- in MBP:CDC42). Für PbHPR1 konnte eine unspezifische Bindung der RNAs beider Stränge gezeigt werden. In der Detektion der Plusstrang-Transkripte scheint eine Bande dominanter zu sein als die übrigen drei, jedoch ist dies ebenso in der eingesetzten RNA und im Überstand zu sehen, so dass vermutlich eine präferenzielle radioaktive Markierung des entsprechenden Oligonukleotides stattfindet. Die Dominanz dieser Bande muss somit als technisches Artefakt angesehen werden und besitzt keine biologische Relevanz (siehe Abbildung 22 C). Das zweite rekombinante Protein MBP:PbHPR2 zeigt keine detektierbare RNA Präzipitation. Die Präzipitation wurde mehrmals mit verschiedenen Puffersystemen und Konzentrationen durchgeführt, es konnte jedoch keine RNA Präzipitation für MBP:PbHPR2 dokumentiert werden.

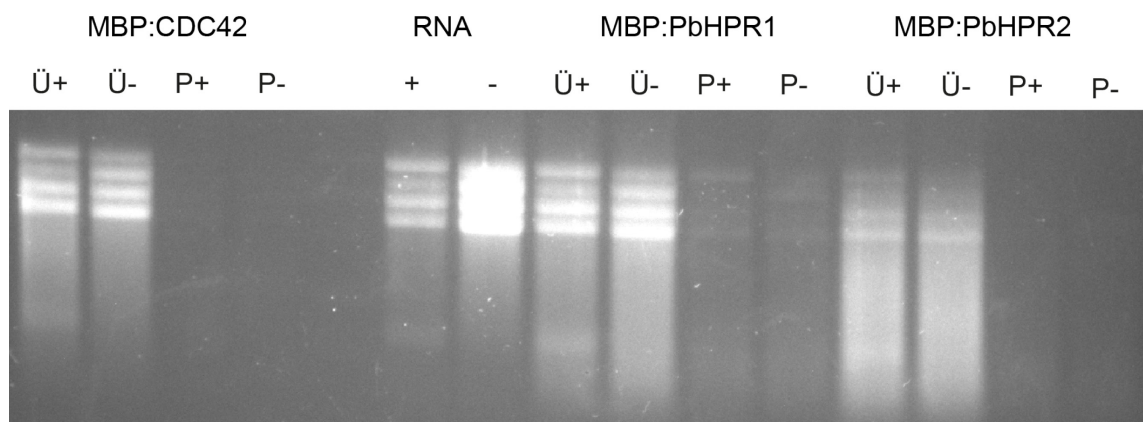


Abbildung 23: Degradation der RNA während der Präzipitation der Fusionsproteine. Die RNA wies während der Präzipitationsexperimente häufig einen variablen Degradationsgrad auf. Während die Anfangs zugegebene RNA intakt ist (Spur RNA + bzw. -) zeigt sich eine leichte Degradation der RNA in den Überständen bei der Inkubation mit MBP:CDC42 und MBP:PbHPR1 (Spuren Ü+ und Ü-). Das Fusionsprotein MBP:PbHPR1 zeigt in der Ethidiumbromid-Färbung eine unspezifische RNA Präzipitation (Spur P+ und P-). Die Überstände des MBP:PbHPR2 Proteins zeigen eine verstärkte Degradation verglichen mit den beiden andern Überständen.

Wie in Abbildung 22 B zu erkennen, findet während der Inkubation der Proteine mit den Transkripten eine Degradation der RNA statt. Um auszuschließen,

dass der Degradationsgrad der RNA einen Einfluss auf die Präzipitation hat, wurden mehrere Replikate angefertigt. Dabei wurde deutlich, dass die Präzipitation der RNA unabhängig vom Degradationsgrad zu sein scheint. Die Degradation mit den Kontrollproteinen variiert partiell, jedoch konnte unabhängig vom Grad der Degradation keine RNA Präzipitation detektiert werden (siehe Abbildung 23). Für PbHPR1 konnte auch in stärker degradierten Proben die unspezifische RNA Präzipitation gezeigt werden (siehe Abbildung 23, Spuren P+ und P- von MBP:PbHPR1). Auffällig bei der Replikation des Experiments ist auch, dass eine stärkere Degradation der RNA im MBP:PbHPR2 Protein erkennbar ist, als bei den anderen zwei Proteinen (Abbildung 23, Spuren Ü+ und Ü- unter MBP:PbHPR2). Einer der Gründe hierfür könnte die C-terminale RAP-Domäne des Proteins sein, der eine potentielle Nukleasefunktion zugeordnet wurde (Boehm et al. 2017). Ob die RAP-Domäne in *in vitro* Experimenten eine erhöhte RNA Degradation verursacht, ist eine interessante Frage für die weitere Analyse dieses Proteins.

### **3.7.4 Rekombinante HPR-Proteine binden *in vitro* keine DNA**

Um für die rekombinanten HPR-Proteine eine generelle Interaktion mit Nukleinsäuren auszuschließen, wurde die gleiche Bindestudie nochmals durchgeführt, dabei wurde jedoch DNA statt RNA eingesetzt. Als Nukleinsäure dienten dabei die in der *in vitro* Transkription verwendeten PCR-Produkte, sodass ein Einfluss der Sequenz minimiert wurde. Auch bei diesem Experiment wurde der Präzipitationserfolg durch Immundetektion, wie oben beschrieben, verfolgt (Abbildung 21). Die Ergebnisse der DNA-Bindestudie sind in Abbildung 24 gezeigt. Die Präzipitation der Fusionsproteine zeigte eine vergleichbare Effizienz wie in der Analyse der RNA Interaktion (siehe Abbildung 24 A). Durch Die Extraktion mit Phenol-Chloroform wurden die Nukleinsäuren aus den jeweiligen Fraktionen isoliert und elektrophoretisch in einem 1 %igen Agarosegel getrennt. Die DNA wurde unter UV-Licht visualisiert und dokumentiert. Auch bei hoher Exposition konnte für keines der Proteine die Präzipitation von DNA detektiert werden (siehe Abbildung 24 B). Während das PbHPR1 eine deutliche Interaktion mit RNA zeigt, konnte für keines der Proteine die Interaktion mit DNA gezeigt werden. Wie in Abbildung 23 gezeigt, kann die Interaktion des Proteins mit RNA bereits in der unspezifischen Ethidiumbromid-Färbung detektiert werden, sodass auf eine radioaktive Detektion



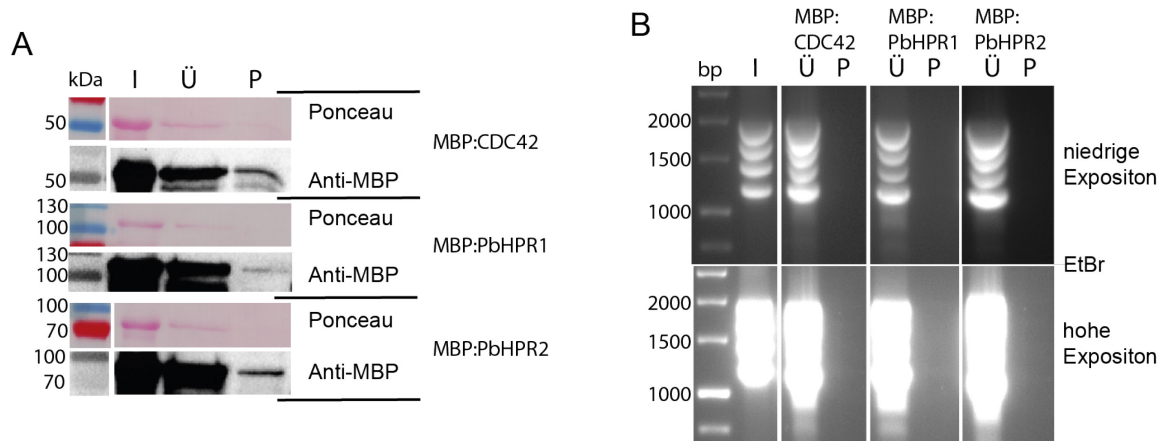


Abbildung 24: Ergebnisse der DNA-Bindestudie mit rekombinanten HPR-Proteinen. Um die Interaktion der Fusionsproteine mit DNA zu überprüfen, wurde eine wie in Abbildung 21 dargestellte Bindestudie mit DNA statt RNA durchgeführt. **A)** Die Immundetektion der Proteinpräzipitation zeigt im eingesetzten Protein (I) und im Überstand (Ü) ein klares Signal für die rekombinanten Proteine. Während der Präzipitationserfolg vergleichbar mit dem in der RNA-Bindestudie ist, wird auch hier nur ein kleiner Teil der Proteine durch die Amylose gebunden (P). In den jeweils oberen Abbildungen ist die unspezifische Proteinfärbung mit Ponceau gezeigt, darunter die Detektion mit spezifischen Antikörpern gegen das MBP-Protein. **B)** Die elektrophoretische Analyse der präzipitierten DNA erfolgte in einem 1 %igen Agarosegel. Die Färbung der DNA erfolgte über die Zugabe von Ethidiumbromid. Während bei der eingesetzte DNA (I), sowie in den Überständen der Proteinpräzipitation (Ü) scharfe Banden für die 4 Fragmente erkennbar sind, ist in den präzipitierten Proteinen keine DNA-Signal erkennbar (P). Auch bei hoher Exposition mit UV-Licht ist kein Signal erkennbar (untere Abbildungen). (teilweise Hillebrand et al. 2018)

der DNA verzichtet wurde. Es lässt sich schlussfolgern, dass PbHPR1 eine wesentlich höhere Bindeaffinität zu RNA als zu DNA hat. Weder für die Kontrolle noch für PbHPR2 konnte eine Interaktion mit entweder RNA oder DNA nachgewiesen werden. Ob dies den experimentellen Parametern geschuldet ist oder ob PbHPR2 auf eine andere Art mit der RNA interagiert, bleibt eine offene Frage.



## 4. Diskussion

Die Expression der Genome von endosymbiontischen Organellen ist abhängig von einer Vielzahl kernkodierter Proteinfaktoren. In dieser Arbeit wurden Proteinfaktoren untersucht, die potentiell im RNA-Metabolismus der Organellen in Parasiten der Gattung *Plasmodium* eine Rolle spielen können. Diese Untersuchung erfolgte mit besonderem Fokus auf die Beteiligung von kernkodierten Proteinfaktoren, die sich durch repetitive helikale Elemente auszeichnen. In höheren Pflanzen ist die Expression der Organellengenome besonders abhängig von den Mitgliedern der PPR-Proteinfamilie (zusammengefasst in z.B. Barkan und Small 2014, Small und Peeters 2000). In *A. thaliana* wurden über 450 der kernkodierten PPR-Proteine beschrieben, von denen die meisten eine experimentell bestätigte oder vorhergesagte Organellenlokalisierung zeigen (Barkan und Small 2014, Colcombet et al. 2013). Die Proteinfamilie ist neben der Prozessierung und Stabilisierung von Transkripten auch an weiteren RNA-Prozessierungsschritten wie der Edierung von Transkripten beteiligt (zusammengefasst in Schmitz-Linneweber und Small 2008, Ruwe et al. 2018, Hammani et al. 2016). Eine sehr wichtige Entdeckung in der Untersuchung von PPR-Proteinen war, dass in vielen Fällen die Bindestelle der Proteine in RNA-Sequenzierungsexperimenten durch die Akkumulation einer kurzen RNA, genannt cosRNA, identifiziert werden kann (z.B. Ruwe und Schmitz-Linneweber 2012, Ruwe et al. 2018, Wang et al. 2017). In *Plasmodium* wurden bisher nur zwei PPR-Proteine identifiziert (Hicks et al. 2018). Diese geringe Zahl an Proteinen in *Plasmodium* scheint kaum ausreichend um die Expression der parasitären Organellengenome zu ermöglichen.

In dieser Arbeit wurde durch die Analyse von Hochdurchsatzsequenzierung eine Population von kurzen, mitochondrialen RNAs identifiziert, die möglicherweise den stabilen Teil einer Proteinbindestelle darstellen könnten. Durch die genomweite Suche nach Proteinen mit repetitiven helikalen Elementen, wurde eine Proteinfamilie identifiziert, deren helikale Elemente strukturell ähnlich zu PPR- und OPR-Motiven sind, jedoch aufgrund von Längenpolymorphismen und Sequenzunterschieden eine neue Proteinfamilie bilden. Im Folgenden sollen die Ergebnisse dieser Arbeit interpretiert und eingeordnet werden.

#### **4.1 Identifikation von cosRNAs in den Organellen von *P. falciparum* als potentielle Interaktionsstellen mit Proteinen**

##### **4.1.1 Die fehlenden cosRNAs im Apicoplasten deuten auf einen PPR-unabhängigen Mechanismus hin**

Die Expression des Apicoplastengenoms ist essentiell für das Überleben und das Wachstum von *Plasmodium* (Ralph et al. 2004, Yeh und DeRisi 2011). Die Gene des Apicoplasten in *Plasmodium* werden polycistronisch transkribiert und anschließend in ihre monocistronische Form prozessiert (Gardner et al. 1991, Nisbet et al. 2016). Während die Expression des Plastidengenoms höherer Pflanzen stark abhängig von PPR-Proteinen ist, scheint dies für *Plasmodium* nicht zuzutreffen, da hier nur 2 PPR-Proteine identifiziert wurden, von denen nur eines eine plastidäre Lokalisation zeigt (Hicks et al. 2018). In dieser Arbeit konnten keine cosRNAs in den intergenischen Bereichen des Apicoplastengenoms detektiert werden, die auf einen ähnlichen Prozessierungsmechanismus wie in höheren Pflanzen hindeuten würden. Es kann allerdings nicht ausgeschlossen werden, dass die apicoplastidären RNAs beispielsweise bei der Erstellung der Sequenzierbibliotheken benachteiligt werden. Während der Untersuchung eines der PPR-Proteine in *P. berghei* wurde ein Prozessierungsmechanismus ähnlich dem *tRNA-punctuation model* in den Mitochondrien der Säugetiere vorgeschlagen (Hicks et al. 2018, Nisbet et al. 2016, Nisbet und McKenzie 2016). In diesem Modell werden bei der Prozessierung von mitochondrialen tRNAs durch RNase P und Z aus polycistronischen Vorläufern, gleichzeitig reife rRNAs und mRNAs generiert (Ojala et al. 1981, zusammengefasst in Taanman 1999). Passend dazu wurde im Genom von *P. falciparum* eine RNase P mit vermutlich plastidärer Lokalisation identifiziert (Lechner et al. 2015). Während im Genom von bspw. humanen Mitochondrien bis auf wenige Ausnahmen die proteinkodierenden und ribosomalen Gene von tRNAs flankiert werden, finden sich im Apicoplastengenom wesentlich mehr Transkripte, die keine flankierende tRNA besitzen (Wilson et al. 1996, Nisbet und McKenzie 2016). Es ist derzeit unklar, ob neben der vermuteten, tRNA-basierten Prozessierung ein weiterer Mechanismus vorliegt, der tRNA unabhängig monocistronische Transkripte generiert. Bei dem im Apicoplasten kodierten Gen *sufB* befindet sich am 5'-Ende ein tRNA-Gen (*trnT*), während am 3'-Ende die

vermutlich proteinkodierenden Gene *orf51* und *orf101* lokalisiert sind (Nisbet et al. 2016). Durch PCR-basierte Methoden konnte gezeigt werden, dass das 5'-Ende des *sufB*-Transkripts mit dem 5'-Ende bzw. 3'-Ende der reifen *trnT*-RNA korrespondieren. Im Gegensatz dazu ist das 3'-Ende von *sufB* stark heterogen mit verschiedenen Transkriptenden innerhalb von *orf51* und *orf101* (Nisbet et al. 2016). Dies könnte darauf hindeuten, dass Transkriptenden im Apicoplasten nur durch tRNAs scharf definiert werden, während tRNA-freie Enden durch die Termination der Transkription oder durch die Stabilität der RNA gegen Exo- bzw. Endonukleasen generiert werden.

#### **4.1.2 In den Mitochondrien von *P. falciparum* akkumulieren cosRNAs an den Transkriptenden**

Die Transkription des mitochondrialen Genoms in *P. falciparum* erfolgt ebenfalls in polycistronischen Einheiten mit einer Länge von bis zu 6 kb, was der vollständigen Genomlänge entspricht (Ji et al. 1996). Eine Prozessierung nach dem tRNA Punktierungsmodell der humanen Mitochondrien ist aufgrund der fehlenden tRNA-Gene in *P. falciparum* unmöglich. Die Hypothese dieser Arbeit ist, dass die Stabilisierung und Prozessierung der mitochondrialen Transkripte in *Plasmodium*, ähnlich wie in pflanzlichen Organellen, durch kernkodierte Proteine vermittelt wird. Ein erster Hinweis auf einen solchen Prozessierungs- bzw. Stabilisierungsmechanismus ist die Identifikation von kurzen RNAs, sogenannten cosRNAs, die an den Enden von mitochondrialen Transkripten akkumulieren (Ruwe et al. 2016, Prikryl et al. 2011, Pfalz et al. 2009). Durch die Analyse von Hochdurchsatzsequenzierungsdaten konnten im Mitochondrium von *P. falciparum* insgesamt 61 distinkte kurze RNAs identifiziert werden. Ein Großteil dieser RNAs konnte bereits charakterisierten Transkripten zugeordnet werden (siehe Tabelle 2), wobei teilweise ein leichtes Abweichen der beschriebenen Enden festgestellt wurde. Neben diesen zugeordneten RNAs wurden weitere kurze RNAs detektiert, die vorher nicht beschrieben wurden bzw. keinem vollständigen Transkript zugeordnet werden konnten. Diese cosRNAs unterscheiden sich von den pflanzlichen cosRNAs durch ihre Positionierung. Während die pflanzlichen cosRNAs meist in den intergenischen Regionen akkumulieren (Ruwe und Schmitz-Linneweber 2012, Ruwe et al. 2016), sind die cosRNAs in *Plasmodium* entweder innerhalb von

definierten Transkripten oder unmittelbar stromauf bzw. stromabwärts von Transkripten zu finden.

Obwohl die Transkription des mitochondrialen Genoms zu einem polycistronischen Primärtranskript führt, konnten für alle drei proteinkodierenden Gene monocistronische Transkripte nachgewiesen werden (Ji et al. 1996). Drei der hier identifizierten cosRNAs korrelieren mit den beschriebenen Enden von proteinkodierenden Genen des mitochondrialen Genoms von *P. falciparum*. Die cosRNA20 überlappt mit dem 5'-Ende der *cob* mRNA (siehe Abbildung 9), die cosRNA24 (5'-Ende *coxIII*) und cosRNA15 (3'-Ende *coxIII*) befinden sich an den Enden der *coxIII* mRNA. Die cosRNAs überlappen mit den bestimmten Transkriptenden in *P. falciparum*. Das Überlappen von cosRNA und Transkriptende wurde ebenfalls für pflanzliche Transkripte beschrieben (Pfalz et al. 2009, Prikryl et al. 2011, Ruwe et al. 2016) und ist ein Hinweis, dass die cosRNAs in *Plasmodium* durch die Bindung von Proteinen entstehen könnten.

Neben der Stabilisierung durch Proteine können cosRNAs in Pflanzen auch unabhängig von Proteinen entstehen. Es konnte gezeigt werden, dass Sekundärstrukturen an den Enden der RNAs ebenfalls stabilisierend wirken können. Auch diese stabilen Sekundärstrukturen können für die Entstehung von cosRNAs verantwortlich sein (Ruwe und Schmitz-Linneweber 2012). Um diese Möglichkeit zu untersuchen, wurde mittels bioinformatischer Programme die potentielle Faltung der cosRNAs untersucht (RNA Faltungsvorhersage mit Mfold<sup>10</sup>). Ein Indikator für die Stabilität einer Sekundärstruktur ist die freie Energie des gefalteten Moleküls. Bei der Analyse pflanzlicher cosRNAs, wurde als Minimalwert für die freie Energie ein Wert von -20 kcal/mol oder niedriger definiert um eine RNA Struktur als stabil zu werten (Ruwe und Schmitz-Linneweber 2012). Nach diesem Kriterium deutet die Vorhersage der RNA-Struktur nicht auf stabile Sekundärstrukturen der cosRNAs in *Plasmodium* hin. Für alle cosRNAs liegt die freie Energie der gefalteten RNA zwischen 0 und -12 kcal/mol. Nur die cosRNA17 könnte mit einer freien Energie von -18 kcal/mol eine stabilere Sekundärstruktur ausbilden. Bei der Suche nach konservierten Sequenzmotiven in den cosRNAs von

---

<sup>10</sup> <http://unafold.rna.albany.edu/?q=mfold>

*P. falciparum*, konnte nur in zwei cosRNAs (cosRNA16 und cosRNA19) am 5'-Ende ein identisches Sequenzmotiv identifiziert werden (TTTATTG). Ansonsten gibt es innerhalb der cosRNAs keine Sequenzhomologien (siehe Anhang Tabelle 5).

#### **4.1.3 Die 5'-Enden von cosRNAs überlappen mit den 5'-Enden von rRNA-Fragmenten**

Für 13 identifizierte cosRNAs wurde das 5'-Ende exakt passend zum 5'-Ende von ribosomalen RNA-Fragmenten kartiert (siehe Tabelle 3). Während die 5'-Enden durch eine extreme Änderung der Sequenzierabdeckung, verglichen mit der vorherigen Genomposition, scharf definiert sind, ist das 3'-Ende häufig weniger genau erkennbar. Die Genomabdeckung nimmt in diesem Bereich stetig ab, ohne einen scharfen Abfall zu zeigen. Während pflanzliche cosRNAs in den intergenischen Bereichen vorliegen, befinden sich die cosRNAs in *P. falciparum* innerhalb der reifen Transkripte. Da die relativ kleinen rRNA Fragmente (zwischen 20 und 200 nt) aus längeren Vorläufern prozessiert werden ist es möglich, dass Proteinfaktoren benötigt werden, um die 5'-Enden während bzw. nach der Prozessierung aus dem polycistronischen Vorläufer zu stabilisieren. Über die Prozesse und Mechanismen durch die die kleinen rRNA Fragmente ins Ribosom inkorporiert werden ist nichts bekannt. Es ist möglich, dass schon während der Transkription eine Bindung der rRNA-Fragmente durch Proteine erfolgt. Wenn eine Kopplung von Transkription und Prozessierung vorliegt wäre eine sofortige Bindung der RNA durch Proteinfaktoren denkbar. Eine Kopplung der Prozessierung mit der mitochondriale Transkription wurde in einem anderen humanen Parasiten beschrieben (Accari und Barth 2015). Die Tatsache, dass die identifizierten cosRNAs bis auf eine Ausnahme nur am 5'-Ende der ribosomalen Fragmente zu finden sind, kann darauf hindeuten, dass die Generierung der 3'-Enden der Fragmente durch einen anderen Mechanismus erfolgt. So konnte für die meisten ribosomalen Fragmente eine Polyadenylierung des 3'-Endes gezeigt werden (Feagin et al. 2012). Auch die Analyse der in dieser Arbeit durchgeführten Sequenzierung deutet auf polyadenylierte Fragmente hin. Generell werden allerdings nur einige wenige Adenine am 3'-Ende von rRNA-Fragmenten detektiert. Dies kann allerdings auch technische Gründe haben, da die bioinformatische Kartierung nicht auf die Detektion längerer Poly-A-Schwänze ausgelegt war. Die

proteinkodierenden Transkripte zeigen ebenfalls Polyadenylierung. Die Länge der Poly-A-Schwänze variiert jedoch stark. Die *cob* mRNA beinhaltet nur einige wenige Adenosine, während für *coxI* bis zu 15 Adenine nachgewiesen wurden (Rehkopf et al. 2000). In menschlichen Mitochondrien ist die Polyadenylierung von Transkripten ein stabilisierendes Element und wird auch für die Generierung von Stopp-Codons benötigt (Gagliardi et al. 2004). Als Hypothese für die Entstehung der cosRNAs innerhalb der rRNA Fragmente könnte also vermutet werden, dass die 5'-Enden der rRNAs durch Proteine stabilisiert werden die hierbei cosRNAs entstehen lassen, während die 3'-Enden durch die Polyadenylierung stabilisiert werden könnten. Dies gilt allerdings nur unter der hypothetischen Annahme, dass die HPR-Proteine bzw. vergleichbare Proteinfaktoren an der Reifung der rRNA-Fragmente beteiligt sind.

Andererseits kann die Entstehung der cosRNAs auch durch die Bindung von ribosomalen Proteinen im Ribosom an die rRNAs erklärt werden. In diesem Fall würden die cosRNAs durch ihre Interaktionen mit Proteinen innerhalb des Ribosoms entstehen. In Feagin et al. 2012 wurden bioinformatisch die rRNA-Fragmente auf die bekannten rRNAs aus *E. coli* angepasst, um potentielle Interaktionspunkte zwischen den einzelnen RNA-Fragmenten zu identifizieren. So formen beispielsweise die beiden Fragmente LSUB und LSUC das für die Funktion des Ribosoms wichtige GTPase Zentrum (Feagin et al. 2012). Durch die Interaktion mehrerer rRNAs könnten sich doppelsträngige RNA-Bereiche ausbilden. Diese Doppelstränge können eine erhöhte Stabilität gegen RNA-Degradation vermitteln und so eine cosRNA verursachen. Allerdings deutet keine der cosRNAs basierend auf Feagin et al. 2012 auf eine solche Situation hin. Zum einen zeigen die cosRNAs keine vollständige Überlappung mit den vorhergesagten Interaktionen, sodass die Ausbildung eines Doppelstranges nur partial erfolgen könnte. Zum anderen würde die Bildung eines Doppelstrangs aus zwei Fragmenten einen längeren Sequenzabschnitt mit komplementären Basen erfordern. Diese müssten folglich auch in der cosRNA enthalten sein. Die Sequenzen der cosRNAs weisen jedoch keine ausgeprägten komplementären Bereiche zueinander. Da es sich jedoch um bioinformatische Vorhersagen für die Ribosomenstruktur handelt, ist eine solche Entstehungsweise nicht vollständig ausgeschlossen, wenn sich die Interaktionen *in vivo* im Ribosom von den Vorhersagen unterscheiden. Um eine abschließende Antwort zu geben, ist nicht nur die Identifikation und Untersuchung der an der RNA

Prozessierung beteiligten Proteine, sondern auch ein tiefergehendes Verständnis der Ribosomenbiogenese im Mitochondrium von *Plasmodium* nötig.

Ein weiterer interessanter Aspekt ist, dass in den polycistronischen Transkripten der Mitochondrien von *Plasmodium* sowohl proteinkodierende als auch ribosomale Fragmente enthalten sind. Wie aus den Sequenzierdaten dieser Arbeit hervorgeht, liegt ein großer Teil der rRNAs in größerer Zahl vor als die proteinkodierenden Transkripte. Es ist also anzunehmen, dass die Stabilisierung der monocistronischen rRNAs regulativ erfolgt, sodass ein Teil der mRNAs abgebaut wird, während die ribosomalen Fragmente stabilisiert werden müssen. Auch hier können Proteinfaktoren eine entscheidende Rolle spielen. Sowohl für OPR- als auch PPR-Proteine wurde diese stabilisierende Wirkung auf Transkripte gezeigt (Pfalz et al. 2009, Meierhoff et al. 2003, Hammani et al. 2014, Wang et al. 2015).

Ebenso wie bei den identifizierten cosRNAs in den proteinkodierenden Transkripten (siehe 4.1.2) sind auch die cosRNAs innerhalb der rRNA-Fragmente im Mitochondrium von *P. falciparum* wesentlich variabler in ihrer Länge (z.B. cosRNA11 = 20 nt, cosRNA10 = 58 nt). Die Länge der cosRNAs in Pflanzen hängt unter anderem von der Größe des korrespondierenden PPR-Proteins ab, da die Stabilisierung des entsprechenden Transkriptes auf der direkten Blockade der Exonukleasen durch das PPR-Protein basiert (Pfalz et al. 2009). Beispielsweise umfasst die minimale Bindestelle des PPR10 (*Z. mays*) nur 17 Nukleotide. Die dabei verursachte cosRNA umfasst 24 Nukleotide (Prikryl et al. 2011). Die HPR-Proteine unterscheiden sich in ihrer molekularen Masse und Anzahl der Repetitionseinheiten nicht wesentlich von den PPR-Proteinen der Pflanzen. Es kann also angenommen werden, dass der vom Protein beschützte Bereich der RNA zwischen HPR- und PPR-Proteinen vergleichbar sein sollte. Wenn die HPR-Proteine an der Entstehung der cosRNAs beteiligt sind, muss sich der Wirkmechanismus von dem der PPR-Proteine unterscheiden. Beispielsweise könnte die Bildung von größeren Proteinkomplexen dazu führen, dass ein längerer Bereich der RNA geschützt werden kann. Auch die Interaktion der HPR-Proteine mit weiteren Proteinen ist eine Möglichkeit um einen längeren Bereich der RNA zu beschützen.

## **4.2 HPR-Proteine – eine neue Familie von helikalen Proteinen im RNA-Metabolismus der Mitochondrien**

### **4.2.1 HPR-Proteine sind strukturell verwandt mit OPR-Proteinen, stellen jedoch eine eigene Familie dar**

Die Expression von Organellengenomen ist stark abhängig von kernkodierten Proteinen, bei denen besonders Proteine mit repetitiven helikalen Elementen eine wichtige Rolle spielen (Barkan und Small 2014, Hammani et al. 2014). Neben den bereits in der Einleitung beschriebenen PPR- und OPR-Proteinen, gibt es weitere Proteinfamilien, die sich durch eine superhelikale Gesamtstruktur auszeichnen und repetitive Sequenzmotive besitzen. Dabei unterscheiden sich die repetitiven Elemente der verschiedenen Proteine nicht nur in der Sequenz, sondern auch in ihrer Länge. Zu diesen Proteinen zählen: Puf-repeat-Proteine (Edwards et al. 2001; 36 AS Länge), HEAT-repeat-Proteine (Groves et al. 1999; 39 AS Länge), TPR-Proteine (Blatch und Lässle 1999; 34 AS) und ARM-repeat-Proteine (Huber et al. 1997; 42 AS) sowie die erwähnten PPR-Proteine (Small und Peeters 2000; 35 AS) und OPR-Proteine (Rahire et al. 2012, Eberhard et al. 2011; 38 AS).

In dieser Arbeit konnte durch bioinformatische Analyse des gesamten Proteoms von *P. falciparum* eine neue Proteinfamilie identifiziert werden, die sich durch ein 37 AS langes, repetitives Element auszeichnet. Zwar zeigen sich innerhalb der ersten Helix stärkere Ähnlichkeiten bei der Sequenz zwischen OPR- und HPR-Motiv, jedoch verschwinden diese Ähnlichkeiten am Ende des ersten helikalen Elements (Position 15 im HPR-Motiv verglichen mit Position 17 des OPR-Motivs, Abbildung 10). Die für die OPR-Proteine definierende PPPEW-Domäne innerhalb des OPR-Motivs fehlt im HPR-Motiv (Eberhard et al. 2011, Rahire et al. 2012). Es besteht eine klare, strukturelle Verwandtschaft zwischen PPR-, OPR- und HPR-Proteinen, jedoch handelt es sich aufgrund der Unterschiede in Sequenz und Länge des Konsensusmotivs bei den HPR-Proteinen vermutlich um eine eigenständige Proteinfamilie.

Unterstützt wird diese Schlussfolgerung durch die Verteilung der HPR-Proteine in den verschiedenen phylogenetischen Gruppen im Vergleich zu der



Verteilung der PPR- und OPR-Proteine (siehe Abbildung 13). Bei der Suche mit dem OPR-Motiv als Suchmotiv wurden nur wenige HPR-Proteine gefunden, während die Suche mit dem HPR-Motiv nur einige wenige OPR-Proteine identifizierte. PPR-Proteine sind in den meisten Eukaryoten zu finden, jedoch sind diese besonders verbreitet in den höheren Pflanzen (Small und Peeters 2000, Barkan und Small 2014, Schmitz-Linneweber und Small 2008). Im Kontrast dazu konnten in *A. thaliana*, einem Vertreter der Embryophyta, keine HPR-Proteine detektiert werden. OPR-Proteine sind bisher nicht phylogenetisch untersucht worden. In dieser Arbeit konnten nur in *C. reinhardtii* mehr OPR- als HPR-Proteine identifiziert werden, ansonsten dominieren die HPR-Proteine in den verschiedenen Taxa. Anhand der hier durchgeführten Verteilungsanalyse ist zu vermuten, dass die OPR-Proteine in *C. reinhardtii* eine für diesen Organismus abgeleitete, spezifische Form darstellen, während die HPR-Proteine eine Proteinfamilie mit breiterer Verteilung sind. Es ist möglich, dass HPR-Proteine eine ähnliche Funktion ausführen wie OPR- und PPR-Proteine und es einem evolutionären Zufall zukommt, dass die Motivlänge zwischen 35, 37 und 38 AS variiert, solange die Hauptfunktion im RNA Metabolismus durch die Motivlänge nicht beeinflusst wird. Um dies aufzulösen und die Herkunft der HPR-Proteine zu klären, ist eine tiefgehende phylogenetische Analyse nötig. Auf einige der identifizierten HPR-Proteine, besonders im Menschen, wird im Folgenden eingegangen.

#### **4.2.2 HPR-Proteine sind vor allem in den meisten Alveolata weit verbreitet**

Die HPR-Proteinfamilie scheint besonders in den Alveolata expandiert zu sein. So wurden neben den Apicomplexa sowohl in Dinoflagellaten als auch in den Chromerida eine große Zahl von HPR-Proteinen detektiert (siehe Abbildung 13), während in den Ciliophora nur 5 HPR-Proteine identifiziert wurden. Eine besonders spannende Beobachtung ist, dass keine HPR-Proteine in *Cryptosporidium hominis* erkennbar waren. In *Cryptosporidium spp.* ist das mitochondriale Genom nicht mehr vorhanden und das mitochondrienähnliche Organelle ist ausschließlich von kernkodierten Faktoren abhängig (Henriquez et al. 2005). Auch der Apicoplast ist in dieser Gruppe reduziert (Zhu et al. 2000). Diese Tatsache liefert einen phylogenetischen Hinweis, dass die HPR-Proteine eine Familie mit Funktion in der organellären Genexpression der Alveolata sein könnte.

#### 4.2.3 Besteht ein Zusammenhang zwischen HPR-Proteinen und fragmentierten rRNA-Genen im mitochondrialen Genom?

Innerhalb der Alveolata ist die Fragmentierung der rRNA-Gene im mitochondrialen Genom weit verbreitet. Neben den meisten Organsimen der Apicomplexa weisen sowohl die Dinoflagellaten, als auch die Chromerida eine ähnliche Fragmentierung des mitochondrialen Genoms auf. Gleichwohl unterscheidet sich der Grad der Fragmentierung. In Apicomplexa sind derzeit 34 einzelne Fragmente beschrieben (Feagin et al. 2012). Das mitochondriale Genom von Dinoflagellaten zeigt einen sehr ähnlichen Gen-Gehalt mit 27 rRNA-Fragmenten und 12 kleinen Transkripten, die den rRNAs aus *Plasmodium* ähneln (Shoguchi et al. 2015). In Vertretern der Chromerida wurden ebenfalls fragmentierte rRNA-Gene identifiziert, die genaue Zahl ist nicht eindeutig geklärt (Flegontov et al. 2015). Innerhalb der Apicomplexa wurden meist mehr als 20 HPR-Proteine detektiert. In den Dinoflagellaten und den Chromerida wurden wesentlich mehr HPR-Proteine identifiziert (*S. minutum* = 73 HPR-Proteine, *C. velia* = 58 HPR-Proteine). Für diese Expansion der HPR-Familie in Dinoflagellaten kann es zwei mögliche Erklärungen geben. Zum einen wurde gezeigt, dass in Dinoflagellaten die Komplexität der mitochondrialen RNA-Prozessierung wesentlich höher ist. So werden mitochondriale Transkripte auf RNA-Ebene ediert und es wurden auch die Transkription von Pseudogenen nachgewiesen (Imanian und Keeling 2007, Waller und Jackson 2009, Nash et al. 2007). Zum anderen wurde im Dinoflagellaten *Karlodinium micrum* gezeigt, dass die reife *coxIII* mRNA durch Transsspleißen aus mindestens zwei Transkripten generiert wird (Jackson et al. 2007, Waller und Jackson 2009). Diese Prozesse benötigen zusätzliche Proteinfaktoren wie etwa HPR-Proteine. Neben dem komplexeren RNA-Metabolismus des Mitochondriums besitzen sowohl Dinoflagellaten als auch Chromerida einen photosynthetisch-aktiven Plastiden, der zusätzliche RNA-Prozessierungsereignisse nötig machen könnte.

Im Kontrast zu den oben genannten Alveolata wurde bei der HPR-Suche in *T. thermophila*, einem Vertreter der Ciliophora, nur 5 HPR-Proteine identifiziert. Eines der Unterscheidungsmerkmale dieses Taxons zu den anderen Taxa der Alveolata ist, dass die Ciliophora keine fragmentierten rRNA-Gene im

mitochondrialen Genom besitzen, sondern zwei durchgängige rRNAs (Slamovits 2014). Die reduzierte Komplexität der mitochondrialen rRNA-Reifung mit nur zwei durchgängigen rRNA-Genen könnte eine Begründung für die Reduktion der HPR-Proteine sein. Während der Fragmentierungsgrad in den Alveolata einzigartig ist, gibt es auch außerhalb dieses Phylums Organismen deren mitochondriales Genom fragmentierte rRNA-Gene zeigen. Einer dieser Vertreter ist *C. reinhardtii*, bei dem die ribosomalen RNA-Gene des Mitochondriums in 4 (kleine Untereinheit) bzw. 8 (große Untereinheit) distinkte Fragmente unterteilt sind (Boer und Gray 1988). Interessanterweise wurden in *C. reinhardtii* 36 HPR-Proteine identifiziert, sodass eine komplexe mitochondriale Genomorganisation auch hier einhergeht mit einer vergleichsweise hohen Zahl von HPR-Proteinen.

In *Plasmodiophora brassicae*, einem einzelligen Parasiten, der in Pflanzen die Krankheit „Kohlhernie“ verursacht und dem heterogenen Taxon Rhizaria zugeordnet wird, wurden 22 HPR-Proteine identifiziert. Während das Genom der Mitochondrien von *P. brassicae* durchgängige rRNA-Gene zeigt, sind die mitochondrialen Transkripte durch eine Vielzahl von Introns unterbrochen (Daval et al. 2018). Es ist also möglich, dass nicht nur die Fragmentierung der mitochondrialen rRNAs mit einer Expansion der HPR-Proteinfamilie einhergeht, sondern auch andere komplexere Prozessierungsereignisse mit einer Verbreiterung der HPR-Proteinfamilie einhergehen. Die Expansion der PPR-Proteine in höheren Pflanzen korreliert ebenfalls mit einem komplexeren RNA-Metabolismus der Organellen, bei dem Transkripte beispielsweise ediert und prozessiert werden müssen (Barkan und Small 2014). Es ist in der Zukunft entscheidend, die genaue Funktion der HPR-Proteine zu ermitteln, um zu entscheiden, ob es einen direkten Zusammenhang zwischen der Expansion der HPR-Proteinfamilie und einem komplexer-werdenden mitochondrialen RNA-Metabolismus gibt.

#### **4.2.4 Die Deletion von HPR-Proteinen führt zu einem dramatischen Fitnessverlust in *P. berghei***

Die Deletion von PPR-Proteinen in höheren Pflanzen verursacht teilweise dramatische Phänotypen. Die Transposon-basierte Deletion des PPR10 Proteins in *Z. mays* führt beispielsweise zu einem weißlichen Phänotyp und die Pflanze ist photosynthetisch inaktiv (Pfalz et al. 2009). Auch die Deletionsmutanten anderer

PPR-Proteine zeigen Entwicklungsdefizite (z.B. Zoschke et al. 2013). Viele Nullmutanten von PPR-Proteinen sind sogar embryo-letal (Schmitz-Linneweber und Small 2008). Die Untersuchung von OPR-Deletionsmutanten zeigten vergleichbare Phänotypen (Kleinknecht et al. 2014, Kleinknecht et al. 2016, Rahire et al. 2012). Die ersten genetischen Untersuchungen der HPR-Proteine in dieser Arbeit deuten ebenfalls auf starke Effekte hin. In dieser Arbeit wurden für 15 HPR-Proteine Deletionsversuche unternommen. Nur für zwei HPR-Proteine konnte der entsprechende Genlocus deletiert werden. Die Analyse dieser Deletionsmutanten steht noch aus, ist aber von essentieller Bedeutung, um eine potentielle Funktion dieser HPR-Proteine zu ermitteln. Für die verbleibenden 13 HPR-Proteine konnten keine lebensfähigen *P. berghei* Parasiten detektiert werden. Bestätigt wird dies durch die genomweiten Deletionsstudien in *P. berghei* (Bushell et al. 2017), *P. falciparum* (Zhang et al. 2018) und *T. gondii* (Sidik et al. 2016). Die Untersuchung von *P. berghei* beinhaltet etwa die Hälfte des gesamten Proteoms. Die Deletion der Proteine erfolgt durch die homologe Rekombination mit Deletionsvektoren der *plasmoGEM*-Bibliothek (Schwach et al. 2015). Von den 22 HPR-Proteinen waren hier 9 in der Analyse enthalten. Von diesen 9 wurden 7 als essentiell klassifiziert. In *P. falciparum* wurden Transposons genutzt, die in AT-reiche Genomabschnitte inserieren. Anschließend wurden die Phänotypen der Parasiten analysiert und durch Sequenzierung die Genomabschnitte identifiziert, in denen das Transposon inseriert war. Von den 22 HPR-Proteinen in *P. falciparum* wurden 20 als essentiell eingestuft. In *T. gondii* wurde das CRISPR/Cas9-System verwendet um die genomweite Analyse durchzuführen. Die Deletionen der HPR-Proteine gehörten zu den stark gegenselektierten Phänotypen. Insgesamt deutet dies darauf hin, dass HPR-Proteine für das Überleben der Parasiten eine große Bedeutung haben.

#### **4.2.5 Welche Faktoren spielen für den Fitnessverlust der Nullmutanten von HPR-Genen eine Rolle?**

Die Deletionsversuche in dieser Arbeit, zusammen mit der genomweiten Untersuchung des essentiellen Proteoms, deuten darauf hin, dass HPR Proteine essentiell sind. Unklar bleibt jedoch, weshalb die HPR-Proteine essentiell sind. Es ist unwahrscheinlich, dass durch die Deletion eines HPR-Proteins ein direkter Fitnessverlust eintritt. Es ist wesentlich wahrscheinlicher, dass durch den Verlust

eines HPR-Proteins sekundäre Effekte auftreten, die die Fitness des Parasiten stark einschränken. Im Gegensatz zu den Mitochondrien anderer Organismen ist in *Plasmodium* die Gewinnung von ATP nur in den Moskitostadien essentiell (Sturm et al. 2015). Einer der wichtigsten Stoffwechselwege im Mitochondrium der Parasiten ist die *de novo* Synthese von Pyrimidinen (Krungkrai und Krungkrai 2016). In Blutstadien von *P. falciparum* konnte gezeigt werden, dass die Elektronentransportkette dem Zweck dient Ubichinon zu regenerieren. Ubichinon wird als Elektronenakzeptor in der Pyrimidinbiosynthese benötigt (Painter et al. 2007). Durch den Elektronentransport entlang der mitochondrialen Membran wird auch das Membranpotential des Mitochondriums aufrechterhalten (Srivastava und Vaidya 1999). Durch das Membranpotential wird der Transport von Metaboliten über die Membran ermöglicht. Zu diesen gehören beispielsweise Aminosäuren und Metall-Ionen (van Dooren et al. 2006). Auch der Import von kernkodierten Proteinen ins Mitochondrium hängt mit dem Membranpotential zusammen (Dudek et al. 2013). Es konnte gezeigt werden, dass die Deletion des ribosomalen L13 Proteins dazu führt, dass die mitochondriale Membran depolarisiert, was schließlich zum Absterben des Parasiten führt (Ke et al. 2018). Es ist zu erkennen, dass die Störung der Biogenese der mitochondrialen Ribosomen zu dramatischen Sekundäreffekten führt. Unter der Annahme, dass die HPR-Proteine an der Reifung bzw. Stabilisierung der rRNA Fragmente beteiligt sind, könnte die Deletion einen ähnlichen Effekt auf die Ribosomen des Mitochondriums haben. Durch den Verlust eines HPR-Proteins würde die Ribosomenbiogenese beeinträchtigt, was in der Folge das Membranpotential beeinflussen würde. Dies könnte schlussendlich den Protein- und Metabolitenimport inhibiert. Dies schränkt die Fitness des Parasiten stark ein.

Es ist auch möglich, dass die HPR-Proteine für die Expression einer oder mehrerer der drei mRNAs des Mitochondriums nötig sind. Die drei Gene des mitochondrialen Genoms kodieren für Untereinheiten der Elektronentransportkette. Eine Störung der Expression eines oder mehrerer dieser Gene führt zur Unterbrechung bzw. Störung der Elektronentransportkette. Die Unterbrechung der Elektronentransportkette ist der Wirkmechanismus des Medikaments Atovaquon, welches die Funktion des Cytochrom bc<sub>1</sub> Komplexes (Komplex III) inhibiert (Fry und Pudney 1992). Die Deletion eines HPR-Proteins, welches für die korrekte Expression eines der drei mitochondrialen Gene wichtig ist, kann also zum

Zusammenbruch der Elektronentransportkette führen und damit die Pyrimidinbiosynthese stören. Dies führt zur Einschränkung der Lebensfähigkeit des Parasiten. Für alle diese Effekte ist auch eine indirekte Interaktion der HPR-Proteine mit anderen Proteinen möglich. Wenn HPR-Proteine nicht direkt an der korrekten Prozessierung bzw. Stabilisierung von RNAs beteiligt sind, sondern für andere Faktoren als Plattform dienen, würde die Deletion eines HPR-Proteins den gleichen Effekt haben. Eine genauere Erklärung für den Phänotyp der HPR-Deletionen kann durch die Identifikation der Funktion der Proteine erfolgen.

#### **4.2.6 HPR-Proteine im Menschen haben eine Funktion im RNA Metabolismus der Mitochondrien**

Während die Funktion der HPR-Proteine in *Plasmodium* durch tiefergehende Analyse noch studiert werden muss, wurden Proteine mit HPR-Motiven im Menschen bereits untersucht. Diese Proteine gehören zur FASTK-Proteinfamilie, die 6 Mitglieder umfasst (FASTK, FASTKD1 – 5). Die Proteine lokalisieren im Mitochondrium und zeichnen sich durch zwei potentielle Kinasedomänen aus (FAST\_1 und FAST\_2, siehe Abbildung 14). Die genaue Funktion dieser Kinasedomänen bzw. ob sie katalytisch aktiv sind, ist derzeit nicht bekannt. Neben diesen Domänen besitzen alle 6 Proteine eine C-terminale RAP-Domäne (Jourdain et al. 2017). Die Untersuchung der einzelnen FASTK-Proteine deutet auf eine Beteiligung am RNA-Metabolismus der humanen Mitochondrien hin (Antonicka und Shoubridge 2015, Wolf und Mootha 2014, Boehm et al. 2017). Das Protein FASTKD5 wurde in RNA-Granula innerhalb des Mitochondriums lokalisiert, wo es an die rRNA der kleinen ribosomalen Untereinheit bindet und für die Ribosomenbiogenese benötigt wird (Antonicka und Shoubridge 2015). Zusätzlich wird FASTKD5 für die Prozessierung der *ATP8/6* und der *COXIII* mRNA benötigt. In Deletionsmutanten von FASTKD5 akkumuliert das unprozessierte Vorläufertranskript, während die prozessierten Formen nicht mehr nachgewiesen werden können (Antonicka und Shoubridge 2015). Auch das Protein FASTKD4 assoziiert mit mitochondrialen mRNAs und beeinflusst die Halbwertszeiten einiger Transkripte (Wolf und Mootha 2014). Möglicherweise ist FASTKD4 auch an der Prozessierung von polycistronischen RNAs beteiligt und dient dem Schutz der Transkripte vor exonukleolytischem Verdau (Boehm et al. 2017).

Alle FASTK-Proteine enthalten im N-terminalen Bereich HPR-Motive. Diese zeigen verglichen mit den HPR-Proteinen aus *Plasmodium* einen höheren Grad der Degeneration, sodass die Erstellung eines FASTK-spezifischen Suchmotives nicht möglich ist. Die Sekundärstrukturvorhersage deutet darauf hin, dass der gesamte Proteinbereich zwischen dem N-terminalen Lokalisationspeptid und der RAP-Domäne durch helikale Strukturelemente dominiert ist. Diese befinden sich auch außerhalb der hier identifizierten HPR-Motive. Aufgrund dieser Sekundärstruktur und den gefundenen HPR-Motiven könnte man die FASTK-Proteine den HPR-Proteinen zurechnen. Dies würde unterstützen, dass HPR-Proteine im RNA-Metabolismus der Mitochondrien eine wichtige Rolle spielen könnten. Theoretisch kann die RNA-Interaktion der FASTK-Proteine auch über die RAP-Domäne erfolgen und den HPR-Motiven eine andere Funktion zukommen. Bei der Untersuchung von FASTKD4 wurde durch Austauschen der RAP-Domäne innerhalb der FASTK-Proteinfamilie festgestellt, dass die RAP-Domäne für die katalytische Funktion der FASTK-Proteine entscheidend ist (Boehm et al. 2017). In diesen Chimären wurde gezeigt, dass die Proteinregion, in der die HPR-Motive identifiziert wurden, für die RNA-Interaktion und die Bildung von RNA-Granula verantwortlich ist (Boehm et al. 2017).

Die funktionelle Studie der FASTK-Proteine deutet darauf hin, dass HPR-Proteine eine ähnliche Funktion erfüllen könnten wie PPR- und OPR-Proteine. Im Vergleich zwischen den HPR-Proteinen in *Plasmodium* und den humanen FASTK-Proteinen zeigen die HPR-Motive in den FASTK-Proteinen einen erhöhten Grad der Degeneration. Dies zeigt sich unter anderem durch höhere e-Werte in der MAST-Suche. Ein Grund dafür kann sein, dass das HPR-Konsensusmotiv nur aus HPR-Motiven generiert wurde, die in Proteinen der Apicomplexa gefunden wurden. Es ist möglich, dass durch evolutionäre Prozesse die Motive in Parasiten und im Menschen sich unabhängig voneinander verändern. Eine mögliche Lösung ist die Erstellung eines HPR-Motivs aus Proteinen, die aus einem breiteren Organismenspektrum stammen. Die HPR-Motive der FASTK-Proteine überlappen meist mit einer der beiden Kinasedomänen (FAST\_1), deren Identifikation in den FASTK-Proteinen auf Homologiestudien basiert (Tian et al. 1995). Derzeit ist nicht klar, ob die FAST\_1 bzw. FAST\_2 Domänen funktional sind (Jourdain et al. 2017). Bei der genaueren Untersuchung der FAST\_1 bzw. FAST\_2 Domänen wurde

festgestellt, dass wichtige Reste im aktiven Zentrum fehlen, die für die Funktion der Kinase essentiell sind (Simarro et al. 2010). Die Überlappung zwischen HPR-Motiven und FAST\_1 Domäne kann ein weiterer Hinweis sein, dass die Kinasedomäne inaktiv ist. Es ist also möglich, dass die HPR-Motive eine ähnliche Struktur bzw. Sequenz beinhalten wie die FAST\_1 und deshalb die Motive sich überdecken. Besonders die FAST\_1 Domäne zeichnet sich durch eine höhere Zahl von Leucinen aus (Jourdain et al. 2017). Nach derzeitigem Stand scheint die nachgewiesene RNA-Interaktion wahrscheinlicher als eine Phosphorylierung von z.B. Proteinen. Letztlich sind aber weitere Untersuchungen nötig um die Funktion der beiden FAST Domänen und der HPR-Motive voneinander abzugrenzen.

### 4.2.7 HPR-Proteine binden *in vitro* unspezifisch RNA

Proteine der PPR-Familie zeichnen sich *in vivo* durch eine sehr spezifische RNA-Erkennung aus und haben daher meist nur eine einzelne, spezifische Bindestelle (Schmitz-Linneweber und Small 2008, Barkan und Small 2014). Die durchgeführten *in vitro* RNA-Interaktionsstudien zeigen jedoch nur eine unspezifische Interaktion von PbHPR1 mit den RNAs des Mitochondriums. Die Frage, die sich stellt, ist nun ob HPR-Proteine auch *in vivo* eine unspezifische RNA-Interaktion zeigen und sich damit funktional von den PPR-Proteinen abheben. Aufgrund der strukturellen Ähnlichkeit zwischen HPR- und PPR-Proteinen würde man vermuten, dass die unspezifische RNA-Interaktion eher den Bedingungen des Experiments geschuldet sind als den physiologischen Eigenschaften des Proteins. So ist es möglich, dass für die spezifische Bindung von PbHPR1 an RNA zusätzliche Proteinfaktoren nötig sind. Beispielsweise wurde in *C. reinhardtii* eine solche Protein-Protein-abhängige RNA-Interaktion gezeigt (Schwarz et al. 2007). Es ist auch möglich, dass die Pufferumgebung der durchgeführten Experimente ursächlich für die unspezifische RNA-Interaktion ist. Für rekombinante Proteine von PPR10 (*Z. mays*) wurde gezeigt, dass die *in vitro* RNA-Interaktion des Proteins stark von der gewählten Pufferumgebung abhängig ist (McDermott et al. 2018). Neben der Pufferumgebung kann vermutlich auch das Verhältnis zwischen RNA und Protein einen Einfluss auf die Interaktion haben.

Während für das MBP keine RNA-Interaktion nachgewiesen wurde, ist es auch möglich, dass die Fusion der beiden Proteine die Bindeeigenschaften



verändert. Eine Möglichkeit dies zu überprüfen ist, die zwischen dem MBP und dem HPR-Protein lokalisierte TEV-Protease Erkennungssequenz zu nutzen. Durch Zugabe der Protease nach der Aufreinigung wäre es möglich, das Fusionsprotein in zwei einzelne Teile zu trennen. Problematisch dabei ist, dass sich HPR-, PPR- und OPR-Proteine aufgrund ihres stark hydrophoben Charakters sich nur schwer als lösliche Proteine in *E. coli* exprimieren lassen (z.B. Rahire et al. 2012). Die Fusion mit dem MBP soll die Löslichkeit des Proteins erhöhen (Kapust und Waugh 1999). Es ist also möglich, dass das Entfernen des MBPs zu unlöslichem HPR-Protein führt und so keine Untersuchung erfolgen kann.

#### **4.2.8 Hypothetisches Modell für die Funktion von HPR-Proteinen im Mitochondrium**

Das Verständnis des RNA-Metabolismus der Mitochondrien in *Plasmodium* ist sehr lückenhaft und auch die Studie der HPR-Proteine steht noch am Anfang. Trotzdem soll im Folgenden anhand der verfügbaren Informationen ein hypothetisches Modell für die Funktion der HPR-Proteine erstellt werden. Anhand dieses Modells sollen zukünftige Untersuchungen und Experimente vorgeschlagen werden, um die Analyse der Funktion der neuen Proteine zu unterstützen. Eine schematische Darstellung des Modells ist in Abbildung 25 gezeigt.

Wie bereits in der Einleitung erwähnt erfolgt die Transkription des Genoms als polycistronisches Transkript durch eine Phagentyp-Polymerase (Li et al. 2001). Die HPR-Proteine binden im 5'-Bereich der rRNA-Fragmente innerhalb des polycistronischen Transkriptes (Abbildung 25, Schritt 1). Die Prozessierung zu monocistronischen Transkripten könnte auf mehrere Arten stattfinden. Bei der Entstehung der pflanzlichen cosRNAs erfolgt die Prozessierung des polycistronischen Transkriptes durch Endonukleasen. Dabei werden kleinere Transkripte generiert. Anschließend werden die Transkriptenden durch die Blockierung von Exonukleasen durch PPR-Proteine definiert (Prikryl et al. 2011, Pfalz et al. 2009). Eine ähnliche Situation könnte sich bei den HPR-Proteine ergeben. Durch bisher nicht bekannte Faktoren könnte das polycistronische Transkript in kleinere Einheiten geschnitten werden. Die HPR-Proteine stabilisieren durch ihre Bindung die 5'-Enden und generieren in Kombination mit Exonukleasen die reifen Transkriptenden. Eine weitere Möglichkeit ist, dass die in fünf HPR-

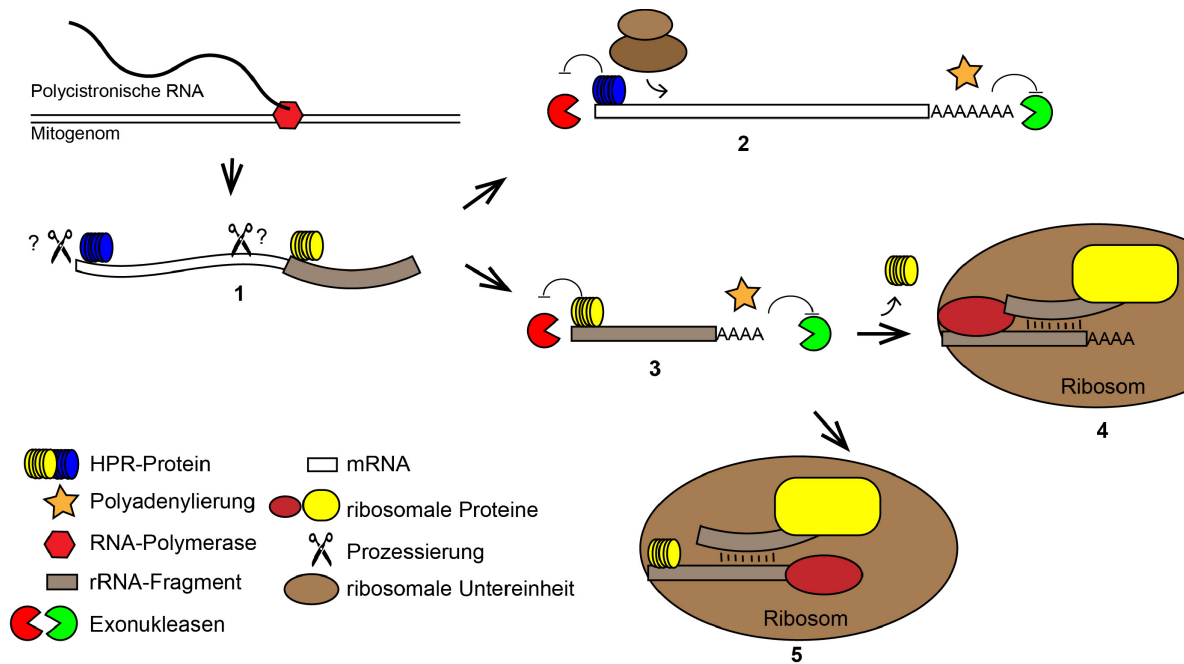


Abbildung 25: Hypothetisches Modell für die Funktion der HPR-Proteine. Beginnend mit der polycistronischen Transkription durch die Polymerase erfolgt in **1)** die Bindung von HPR-Proteinen an spezifische Sequenzen im polycistronischen Vorläufer. Wie die Prozessierung der polycistronischen RNA zu kleineren Einheiten erfolgt ist unbekannt. Eine Möglichkeit ist, dass Proteine mit endonukleolytischer Aktivität das primäre Transkript schneiden, die reifen Enden der Transkripte werden anschließend durch Exonukleasen oder durch die HPR-Proteine selbst definiert. **2)** Die mRNAs werden durch HPR-Proteine am 5'-Ende stabilisiert und durch die Polyadenylierung des 3'-Endes werden ebenfalls Exonukleasen blockiert. Möglicherweise spielen die HPR-Proteine auch eine Rolle bei der Translationsregulation, wenn die HPR-Bindestelle nicht das Startcodon verdeckt. **3)** rRNA-Fragmente werden wie mRNAs durch HPR-Proteine am 5'-Ende vor dem Abbau geschützt. Die polyadenylierten 3'-Enden schützen ebenfalls vor exonukleolytischem Verdau. **4)** Die rRNA wird in das Ribosom eingebaut, dabei wird das Fragment durch ribosomale Proteine gebunden und andere rRNA-Fragmente interagieren über Wasserstoffbrücken miteinander. Das HPR-Protein löst sich dabei von der RNA und ist nicht Teil des Ribosoms. **5)** Alternativ zu 4) wird hier das rRNA-Fragment in das Ribosom integriert, jedoch bleibt das HPR-Protein an der RNA gebunden. Potentiell ist das HPR-Protein als Teil des aktiven Ribosoms auch an der Translationsregulation beteiligt. Für Details siehe Text.

Proteinen in *P. falciparum* enthaltene RAP-Domäne die Prozessierung durchführt und keine weiteren Prozessierungsfaktoren nötig sind. Es ist auch möglich, dass die Prozessierung durch andere Proteine erfolgt, die eine endonukleolytische Aktivität besitzen und mit den HPR-Proteinen interagieren. Um die Aktivität der

RAP-Domäne zu untersuchen, wäre beispielsweise die Inaktivierung der RAP-Domäne durch gerichtete Mutagenese möglich. Während der Untersuchung der FASTK-Proteine wurden mögliche aktive Reste der RAP-Domäne identifiziert, die für eine solche Mutagenese genutzt werden könnten (Boehm et al. 2017). Um mögliche Interaktionspartner der HPR-Proteine zu identifizieren, wäre beispielsweise eine Immunpräzipitation der HPR-Proteine mit spezifischen Antikörpern durchzuführen. Die Identifikation der kopräzipitierten Proteine könnte durch massenspektroskopische Methoden erfolgen. Alternativ könnten die 5'-Enden alleine durch die Blockierung von 5'- nach 3'-Exonukleasen verursacht werden, wie es für die Entstehung von chloroplastidären 5'-Enden von RNAs durch PPR-Proteine gezeigt worden ist.

Nach der Prozessierung findet am 3'-Ende sowohl für proteinkodierende Transkripte als auch für rRNA-Fragmente eine Polyadenylierung durch einen unbekannten Proteinfaktor statt (Feagin et al. 2012, Rehkopf et al. 2000). Das am 5'-Ende gebundene HPR-Protein könnte also das Transkript vor Exonukleasen beschützen, ähnlich wie für PPR-Proteine beschrieben (Prikryl et al. 2011, Meierhoff et al. 2003, Pfalz et al. 2009). Dagegen könnte das 3'-Ende entweder durch ein zusätzliches Protein oder durch die Polyadenylierung stabilisiert werden. Gleichzeitig kann die Translation durch die Ribosomen stattfinden (Abbildung 25 Schritt 2). Für einige PPR-Proteine ist eine Aktivierung der Translation nachgewiesen, möglicherweise hätten die am 5'-Ende gebundenen HPR-Proteine also nicht nur eine stabilisierende Funktion, sondern sind auch für die Translation der mRNA nötig (Manavski et al. 2012, Haïli et al. 2016). Eine experimentelle Untersuchung an dieser Stelle ist schwierig, da die Deletion von HPR-Proteinen nicht möglich ist. Eine Möglichkeit dies trotzdem zu untersuchen wäre, die Expressionslevel des HPR-Proteins durch einen Promotoraustausch zu modifizieren und anschließend sowohl die Transkript- als auch die Proteinmengen der mitochondrialen Gene zu untersuchen. Da die einsetzbare Menge von *P. berghei* aufgrund der Kultivierung in Mäusen limitiert ist, wäre alternativ eine Untersuchung in Organsimen wie *T. gondii* möglich, da hier das Kulturvolumen und damit die Parasitenmenge besser skaliert werden kann.

Die Stabilisierung der rRNA-Fragmente durch HPR-Proteine könnte genauso erfolgen wie für die mRNAs (Abbildung 25 Schritt 3). Unter der Annahme, dass jede cosRNA durch ein HPR-Protein verursacht wird, deuten die Daten in dieser Arbeit auf eine präferenzielle Bindung der 5'-Enden der rRNA-Fragmente hin. Nach der Prozessierung und Stabilisierung der rRNA-Fragmente müssen die Ribosomen assembliert werden. Hier gibt es zwei mögliche Wege wie die HPR-Proteine arbeiten könnten. Im ersten Weg (Abbildung 25 Schritt 4) werden die rRNA-Fragmente in die entsprechende Untereinheit des Ribosoms inkorporiert und die HPR-Proteine dabei von der RNA entfernt. Die Stabilisierung erfolgt dann durch Interaktionen im Ribosom mit ribosomalen Proteinen oder durch RNA-RNA-Interaktion. Das HPR-Protein ist dabei nicht Teil des Ribosoms. Für das humane FASTKD2 wurde in Mutanten ein Defekt in der Assemblierung der großen ribosomalen Untereinheit festgestellt (Antonicka und Shoubridge 2015). Auch die Deletion von FASTKD5 führte zu einer globalen Reduktion der mitochondrialen Translation. Gleichzeitig wurden die Zellbestandteile durch Zentrifugation in Dichtegradienten analysiert. Es konnte gezeigt werden, dass FASTKD2 und FASTKD5 in Dichtegradienten nicht mit den Ribosomen zusammen wandern und somit nicht Teil des aktiven Ribosoms sind (Antonicka und Shoubridge 2015). Der zweite vorgeschlagene Weg führt dazu, dass die HPR-Proteine bei der Assemblierung des Ribosoms in die ribosomale Untereinheit inkorporiert werden und dabei an die RNA gebunden bleiben (Abbildung 25 Schritt 5). Für einen Teil der PPR-Proteine in Organsimen der Gattung *Trypanosoma* wurde gezeigt, dass die Proteine Teil des aktiven Ribosoms sind und dabei nicht nur für die Stabilisierung der RNA verantwortlich sind, sondern die Translation beeinflussen (Aphasizheva et al. 2015). Auch in den Ribosomen von pflanzlichen Mitochondrien konnten zahlreiche PPR-Proteine identifiziert werden (Waltz et al. 2019). In diesem Falle käme den HPR-Proteinen also auch eine regulatorische Funktion im Ribosom zu. Eine Möglichkeit um zwischen diesen beiden Wegen zu unterscheiden wäre die Analyse von Zelllysaten von *Plasmodium* in einem Saccharose-Gradienten. Durch beispielsweise Immundetektion müsste die Wanderung der Ribosomen und der HPR-Proteine im Gradienten verfolgt werden. Für den Fall, dass die HPR-Proteine Teil des Ribosoms sind, würde man eine Komigration der beiden beobachten. Sollten HPR-Proteine jedoch frei oder in kleineren Komplexen mit einzelnen rRNAs

vorliegen, würden die wesentlich schwereren Ribosomen deutlich tiefer im Gradienten vorliegen, als die HPR-Proteine.

Die Überprüfung dieses hypothetischen Modells erfordert vor allem die Identifikation der Bindestellen der HPR-Proteine. Diese Identifikation kann durch verschiedenen Strategien erfolgen. Aufgrund des größtenteils essentiellen Phänotyps bei HPR-Deletionen ist die *in vitro* Untersuchung von rekombinanten HPR-Proteinen eine attraktive Möglichkeit. Jedoch muss durch Optimierung die spezifische RNA-Bindung *in vitro* erreicht werden. Die *in vitro* identifizierten Bindestellen müssen anschließend durch *in vivo* Methoden bestätigt werden. Für die *in vivo* Untersuchung ist ein wesentliches Problem das relativ niedrige Expressionslevel der meisten HPR-Proteine. Die Überexpression der HPR-Proteine ist eine mögliche Lösung hierfür. Eine Alternative für die *in vivo* Untersuchung der Bindestelle kann durch die genetische Markierung von HPR-Proteinen mit RNA-modifizierenden Enzymen erfolgen. So konnte für PUF-Proteine *in vivo* ein Fusionsprotein mit einer Poly-Uridin-Polymerase (PUP) hergestellt werden, die selektiv die vom PUF-Protein gebundenen RNAs polyuridiniert (Lapointe et al. 2015). In der anschließenden Analyse des Transkriptoms durch RNA-Sequenzierung konnten multiple RNAs anhand der Poly-U-Schwänze als Interaktionspartner des PUF-Proteins identifiziert werden (Lapointe et al. 2015). Ein ähnlicher Ansatz wird als *TRIBES* (*targets of RNA-binding proteins identified by editing*) bezeichnet, bei dem das RNA-Bindeprotein mit einem RNA-Edierungsenzym fusioniert wird (McMahon et al. 2016). Bei dem Enzym handelt es sich um eine ADAR (*adenosine deaminase acting on RNA*), die die Deaminierung einer Adeninbase in einer doppelsträngigen RNA zu Inosin katalysiert (Nishikura 2015). Bei der angeschlossenen Analyse der RNA durch RNA-Sequenzierung kann diese Edierung verfolgt werden und lässt so Rückschlüsse auf die Bindestelle des RNA-Bindeproteins zu (McMahon et al. 2016). In jedem Fall ist die Identifikation der Zielsequenzen von HPR-Proteinen eines der drängendsten Probleme in diesem Feld, dessen Lösung weitere funktionelle Analysen für die Rolle von HPR-Proteinen in der mitochondrialen RNA Prozessierung erst möglich machen wird.

#### 4.2.9 Eignen sich die HPR-Proteine für die Entwicklung neuer Medikamente gegen Malaria?

Im Kampf gegen Malaria stellt die Entstehung von Resistenzen gegen Medikamente eine große Herausforderung dar. Medikamente wie etwa Chloroquin sind aufgrund der entstandenen Resistenzen heute kaum noch in Verwendung (Payne 1987, Wellems und Plowe 2001). Durch die grundlegende Forschung können neue zelluläre Funktionen und Mechanismen in *Plasmodium* identifiziert werden, die einen neuen Ansatzpunkt für die gezielte Entwicklung von Medikamenten bieten. Hierbei können sowohl der Apicoplast, als auch das Mitochondrium ein potentiell Ziel darstellen (Mukherjee und Sadhukhan 2016, Ralph et al. 2001, Nixon et al. 2013, Goodman et al. 2017). Die interessante Frage ist, ob die in dieser Arbeit neu beschriebenen HPR-Proteine als Ansatzpunkt für eine Medikamentenentwicklung dienen können. Die meisten theoretischen Ansätze, die darauf abzielen Protein-Protein bzw. Protein-RNA Interaktionen zu stören, basieren auf dem Einsatz von niedermolekularen Substanzen, die kompetitiv um die Interaktionsoberfläche der Proteine mit RNA konkurrieren und so eine Interaktion blockieren sollen (Aboul-ela 2009, Hermann 2000, Connelly et al. 2016). Ein solcher Ansatz ist in Bezug auf die HPR-Proteine problematisch. Da zu vermuten ist, dass die HPR-Proteine einen ähnliche RNA-Interaktionsmechanismus wie die FASTK-Proteine besitzen, ist auch bei einer spezifischen Wirkung der Inhibition eine größere Anzahl von ungewollten Interaktionen mit den menschlichen Proteinen zu erwarten. Es konnte gezeigt werden, dass die Deletion von FASTK-Proteinen schwerwiegende Erkrankungen auslösen können (z.B. Ghezzi et al. 2008). Da die FASTK-Proteine ebenfalls HPR-Motive enthalten, ist hier eine Kreuzwirkung absehbar.

Eine weitere Erschwernis für die medizinische Nutzung der HPR-Proteine ist, dass neben den HPR-Motiven nur wenige katalytisch aktive Domänen identifiziert werden konnten. Die einzige aktive Domäne ist derzeit die RAP-Domäne, die ebenfalls in den FASTK-Proteinen enthalten ist, sodass auch hier eine katalytische Inhibition ungewollte Nebeneffekte produzieren würde. Für PPR-Proteine wurden weitere aktive, c-terminale Domänen beschrieben, die beispielsweise in der Edierung von Transkripten beteiligt sind (z.B. Sugita und Ichinose 2018). Außer der

RAP-Domäne sind für OPR-Proteine keine weiteren konservierten Domänen beschrieben. Es ist zwar möglich, dass die HPR-Proteine weitere Domänen enthalten, die für eine medizinische Anwendung genutzt werden könnten, jedoch wurde bis jetzt keine Domäne dieser Art identifiziert. Zuletzt ist es möglich, nicht die direkte Interaktion der HPR-Proteine zum Ziel zu machen, sondern die nachfolgenden Prozesse, allerdings ist dafür der Wissensstand über den RNA-Metabolismus und die Proteinbiosynthese der Mitochondrien in *Plasmodium* zu gering. Derzeit sind HPR-Proteine vor allem interessant um die grundlegenden Prozesse des Mitochondriums zu verstehen, eine medizinische Nutzung scheint nach derzeitigem Stand eher unwahrscheinlich.

## Literaturverzeichnis

- Aboul-ela F. 2009. Strategies for the design of RNA-binding small molecules. *Future Medicinal Chemistry* 2: 93-119, <https://doi.org/10.4155/fmc.09.149>
- Accari J. und Barth C. 2015. Transcription and processing of mitochondrial RNA in the human pathogen *Acanthamoeba castellanii*. *Mitochondrion* 23: 25-31, <https://doi.org/10.1016/j.mito.2015.05.003>
- Agrawal S., Chung D.-W. D., Ponts N., van Dooren G. G., Prudhomme J., Brooks C. F., Rodrigues E. M., Tan J. C., Ferdig M. T., Striepen B. und Le Roch K. G. 2013. An Apicoplast Localized Ubiquitylation System Is Required for the Import of Nuclear-encoded Plastid Proteins. *PLOS Pathogens* 9, <https://doi.org/10.1371/journal.ppat.1003426>
- Aikawa M., Hepler P. K., Huff C. G. und Sprinz H. 1966. The feeding mechanism of avian malarial parasites. *The Journal of cell biology* 28: 355-73, <https://doi.org/10.1083/jcb.28.2.355>
- Antonicka H. und Shoubbridge Eric A. 2015. Mitochondrial RNA Granules Are Centers for Posttranscriptional RNA Processing and Ribosome Biogenesis. *Cell Reports* 10: 920-32, <https://doi.org/10.1016/j.celrep.2015.01.030>
- Antony H. A. und Parija S. C. 2016. Antimalarial drug resistance: An overview. *Tropical parasitology* 6: 30-41, <https://doi.org/10.4103/2229-5070.175081>
- Aphasizheva I., Maslov D., Wang X., Huang L. und Aphasizhev R. 2011. Pentatricopeptide Repeat Proteins Stimulate mRNA Adenylation/Uridylation to Activate Mitochondrial Translation in Trypanosomes. *Molecular Cell* 42: 106-17, <https://doi.org/10.1016/j.molcel.2011.02.021>
- Aphasizheva I., Maslov D. A., Qian Y., Huang L., Wang Q., Costello C. E. und Aphasizhev R. 2015. Ribosome-associated pentatricopeptide repeat proteins function as translational activators in mitochondria of trypanosomes. *Molecular Microbiology* 99: 1043-58, <https://doi.org/10.1111/mmi.13287>
- Auchincloss A. H., Zerges W., Perron K., Girard-Bascou J. und Rochaix J.-D. 2002. Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast *psbC* mRNA in *Chlamydomonas reinhardtii*. *The Journal of cell biology* 157: 953, <https://doi.org/10.1083/jcb.200201060>
- Aurrecoechea C., et al. 2009. PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research* 37: D539-D43, <https://doi.org/10.1093/nar/gkn814>
- Bailey T. L., Boden M., Buske F. A., Frith M., Grant C. E., Clementi L., Ren J., Li W. W. und Noble W. S. 2009. MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research* 37: W202-W08, <https://doi.org/10.1093/nar/gkp335>



- Bannister L. H., Hopkins J. M., Fowler R. E., Krishna S. und Mitchell G. H. 2000. A Brief Illustrated Guide to the Ultrastructure of *Plasmodium falciparum* Asexual Blood Stages. *Parasitology Today* 16: 427-33, [https://doi.org/10.1016/S0169-4758\(00\)01755-5](https://doi.org/10.1016/S0169-4758(00)01755-5)
- Barkan A., Rojas M., Fujii S., Yap A., Chong Y. S., Bond C. S. und Small I. 2012. A Combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. *PLOS Genetics* 8: e1002910, <https://doi.org/10.1371/journal.pgen.1002910>
- Barkan A. und Small I. 2014. Pentatricopeptide Repeat Proteins in Plants. *Annual Review of Plant Biology* 65: 415-42, <https://doi.org/10.1146/annurev-arplant-050213-040159>
- Behal R. H., Buxton D. B., Robertson J. G. und Olson M. S. 1993. Regulation of the Pyruvate Dehydrogenase Multienzyme Complex. *Annual Review of Nutrition* 13: 497-520, <https://doi.org/10.1146/annurev.nu.13.070193.002433>
- Bender A., van Dooren G. G., Ralph S. A., McFadden G. I. und Schneider G. 2003. Properties and prediction of mitochondrial transit peptides from *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 132: 59-66, <https://doi.org/10.1016/j.molbiopara.2003.07.001>
- Bertani G. 1951. Studies on Lysogenesis I. *Journal of Bacteriology* 62: 293, PMID:14888646
- Bhatt S., et al. 2015. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* 526: 207, <https://doi.org/10.1038/nature15535>
- Blatch G. L. und Lässle M. 1999. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays* 21: 932-39, [https://doi.org/10.1002/\(SICI\)1521-1878\(199911\)21:11<932::AID-BIES5>3.0.CO;2-N](https://doi.org/10.1002/(SICI)1521-1878(199911)21:11<932::AID-BIES5>3.0.CO;2-N)
- Boehm E., Zaganelli S., Maundrell K., Jourdain A. A., Thore S. und Martinou J.-C. 2017. FASTKD1 and FASTKD4 have opposite effects on expression of specific mitochondrial RNAs, depending upon their endonuclease-like RAP domain. *Nucleic Acids Research* 45: 6135-46, <https://doi.org/10.1093/nar/gkx164>
- Boer P. H. und Gray M. W. 1988. Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. *Cell* 55: 399-411, [https://doi.org/10.1016/0092-8674\(88\)90026-8](https://doi.org/10.1016/0092-8674(88)90026-8)
- Bryant C., Voller A. und Smith M. J. H. 1964. The Incorporation of Radioactivity from [<sup>14</sup>C] Glucose into the Soluble Metabolic Intermediates of Malaria Parasites\*. *The American journal of tropical medicine and hygiene* 13: 515-19, <https://doi.org/10.4269/ajtmh.1964.13.515>

- Bushell E., Gomes A. R., Sanderson T., Anar B., Girling G., Herd C., Metcalf T., Modrzynska K., Schwach F., Martin R. E., Mather M. W., McFadden G. I., Parts L., Rutledge G. G., Vaidya A. B., Wengelnik K., Rayner J. C. und Billker O. 2017. Functional Profiling of a *Plasmodium* Genome Reveals an Abundance of Essential Genes. *Cell* 170: 260-72.e8, <https://doi.org/10.1016/j.cell.2017.06.030>
- Charan M., Choudhary H. H., Singh N., Sadik M., Siddiqi M. I., Mishra S. und Habib S. 2017. [Fe–S] cluster assembly in the apicoplast and its indispensability in mosquito stages of the malaria parasite. *The FEBS Journal* 284: 2629-48, <https://doi.org/10.1111/febs.14159>
- Chomczynski P. und Sacchi N. 2006. The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction: twenty-something years on. *Nature Protocols* 1: 581, <https://doi.org/10.1038/nprot.2006.83>
- Claros M. G. und Vincens P. 1996. Computational Method to Predict Mitochondrially Imported Proteins and their Targeting Sequences. *European Journal of Biochemistry* 241: 779-86, <https://doi.org/10.1111/j.1432-1033.1996.00779.x>
- Colcombet J., Lopez-Obando M., Heurtevin L., Bernard C., Martin K., Berthomé R. und Lurin C. 2013. Systematic study of subcellular localization of *Arabidopsis* PPR proteins confirms a massive targeting to organelles. *RNA Biology* 10: 1557-75, <https://doi.org/10.4161/rna.26128>
- Connelly Colleen M., Moon Michelle H. und Schneekloth John S. 2016. The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chemical Biology* 23: 1077-90, <https://doi.org/10.1016/j.chembiol.2016.05.021>
- Cowman A. F., Healer J., Marapana D. und Marsh K. 2016. Malaria: Biology and Disease. *Cell* 167: 610-24, <https://doi.org/10.1016/j.cell.2016.07.055>
- Dahl E. L. und Rosenthal P. J. 2007. Multiple Antibiotics Exert Delayed Effects against the *Plasmodium falciparum* Apicoplast. *Antimicrobial Agents and Chemotherapy* 51: 3485, <https://doi.org/10.1128/AAC.00527-07>
- Dahl E. L., Shock J. L., Shenai B. R., Gut J., DeRisi J. L. und Rosenthal P. J. 2006. Tetracyclines Specifically Target the Apicoplast of the Malaria Parasite *Plasmodium falciparum*. *Antimicrobial Agents and Chemotherapy* 50: 3124, <https://doi.org/10.1128/AAC.00394-06>
- Daval S., Belcour A., Gazengel K., Legrand L., Gouzy J., Cottret L., Lebreton L., Aigu Y., Mougél C. und Manzanares-Dauleux M. J. 2018. Computational analysis of the *Plasmodiophora brassicae* genome: mitochondrial sequence description and metabolic pathway database design. *Genomics* <https://doi.org/10.1016/j.ygeno.2018.11.013>

- Davies S. M. K., Lopez Sanchez M. I. G., Narsai R., Shearwood A.-M. J., Razif M. F. M., Small I. D., Whelan J., Rackham O. und Filipovska A. 2012. MRPS27 is a pentatricopeptide repeat domain protein required for the translation of mitochondrially encoded proteins. *FEBS Letters* 586: 3555-61, <https://doi.org/10.1016/j.febslet.2012.07.043>
- Desai U. J. und Pfaffle P. K. 1995. Single-step purification of a thermostable DNA polymerase expressed in *Escherichia coli*. *Biotechniques* 19: 780-4, PMID:8588916
- Douglas R. G., Amino R., Sinnis P. und Frischknecht F. 2015. Active migration and passive transport of malaria parasites. *Trends in Parasitology* 31: 357-62, <https://doi.org/10.1016/j.pt.2015.04.010>
- Dubremetz J. F., Garcia-Réguet N., Conseil V. und Fourmaux M. N. 1998. Apical organelles and host-cell invasion by Apicomplexa. *International Journal for Parasitology* 28: 1007-13, [https://doi.org/10.1016/S0020-7519\(98\)00076-9](https://doi.org/10.1016/S0020-7519(98)00076-9)
- Dudek J., Rehling P. und van der Laan M. 2013. Mitochondrial protein import: Common principles and physiological networks. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1833: 274-85, <https://doi.org/10.1016/j.bbamcr.2012.05.028>
- Eberhard S., Loiselay C., Drapier D., Bujaldon S., Girard-Bascou J., Kuras R., Choquet Y. und Wollman F.-A. 2011. Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atpA* transcripts in *Chlamydomonas reinhardtii* chloroplasts. *The Plant Journal* 67: 1055-66, <https://doi.org/10.1111/j.1365-313X.2011.04657.x>
- Edwards T. A., Pyle S. E., Wharton R. P. und Aggarwal A. K. 2001. Structure of Pumilio Reveals Similarity between RNA and Peptide Binding Motifs. *Cell* 105: 281-89, [https://doi.org/10.1016/S0092-8674\(01\)00318-X](https://doi.org/10.1016/S0092-8674(01)00318-X)
- Elliott D. A., McIntosh M. T., Hosgood H. D., 3rd, Chen S., Zhang G., Baeovova P. und Joiner K. A. 2008. Four distinct pathways of hemoglobin uptake in the malaria parasite *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences of the United States of America* 105: 2463-68, <https://doi.org/10.1073/pnas.0711067105>
- Emanuelsson O., Nielsen H., Brunak S. und von Heijne G. 2000. Predicting Subcellular Localization of Proteins Based on their N-terminal Amino Acid Sequence. *Journal of Molecular Biology* 300: 1005-16, <https://doi.org/10.1006/jmbi.2000.3903>
- Esseiva A. C., Naguleswaran A., Hemphill A. und Schneider A. 2004. Mitochondrial tRNA Import in *Toxoplasma gondii*. *Journal of Biological Chemistry* 279: 42363-68,
- Fairhurst R. M., Nayyar G. M. L., Breman J. G., Hallett R., Vennerstrom J. L., Duong S., Ringwald P., Wellems T. E., Plowe C. V. und Dondorp A. M. 2012.

- Artemisinin-resistant malaria: research challenges, opportunities, and public health implications. *The American journal of tropical medicine and hygiene* 87: 231-41, <https://doi.org/10.4269/ajtmh.2012.12-0025>
- Feagin J. E. 1992. The 6-kb element of *Plasmodium falciparum* encodes mitochondrial cytochrome genes. *Molecular and Biochemical Parasitology* 52: 145-48, [https://doi.org/10.1016/0166-6851\(92\)90046-M](https://doi.org/10.1016/0166-6851(92)90046-M)
- Feagin J. E. und Drew M. E. 1995. *Plasmodium falciparum*: Alterations in Organelle Transcript Abundance during the Erythrocytic Cycle. *Experimental Parasitology* 80: 430-40, <https://doi.org/10.1006/expr.1995.1055>
- Feagin J. E., Harrell M. I., Lee J. C., Coe K. J., Sands B. H., Cannone J. J., Tami G., Schnare M. N. und Gutell R. R. 2012. The Fragmented Mitochondrial Ribosomal RNAs of *Plasmodium falciparum*. *PLOS ONE* 7: e38320, <https://doi.org/10.1371/journal.pone.0038320>
- Feagin J. E., Mericle B. L., Werner E. und Morris M. 1997. Identification of additional rRNA fragments encoded by the *Plasmodium falciparum* 6 kb element. *Nucleic Acids Research* 25: 438-46, <https://doi.org/10.1093/nar/25.2.438>
- Feagin J. E., Werner E., Gardner M. J., Williamson D. H. und Wilson R. J. M. 1992. Homologies between the contiguous and fragmented rRNAs of the two *Plasmodium falciparum* extrachromosomal DNAs are limited to core sequences. *Nucleic Acids Research* 20: 879-87, <https://doi.org/10.1093/nar/20.4.879>
- Flegontov P., Michálek J., Janouškovec J., Lai D.-H., Jirků M., Hajdušková E., Tomčala A., Otto T. D., Keeling P. J., Pain A., Oborník M. und Lukeš J. 2015. Divergent Mitochondrial Respiratory Chains in Phototrophic Relatives of Apicomplexan Parasites. *Molecular Biology and Evolution* 32: 1115-31, <https://doi.org/10.1093/molbev/msv021>
- Foth B. J., Ralph S. A., Tonkin C. J., Struck N. S., Fraunholz M., Roos D. S., Cowman A. F. und McFadden G. I. 2003. Dissecting Apicoplast Targeting in the Malaria Parasite *Plasmodium falciparum*. *Science* 299: 705, <https://doi.org/10.1126/science.1078599>
- Foth B. J., Stimmler L. M., Handman E., Crabb B. S., Hodder A. N. und McFadden G. I. 2004. The malaria parasite *Plasmodium falciparum* has only one pyruvate dehydrogenase complex, which is located in the apicoplast. *Molecular Microbiology* 55: 39-53, <https://doi.org/10.1111/j.1365-2958.2004.04407.x>
- Freese N. H., Norris D. C. und Loraine A. E. 2016. Integrated genome browser: visual analytics platform for genomics. *Bioinformatics* 32: 2089-95, <https://doi.org/10.1093/bioinformatics/btw069>
- Fry M. und Pudney M. 1992. Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4-chlorophenyl) cyclohexyl]-3-hydroxy-

- 1,4-naphthoquinone (566C80). *Biochemical Pharmacology* 43: 1545-53, [https://doi.org/10.1016/0006-2952\(92\)90213-3](https://doi.org/10.1016/0006-2952(92)90213-3)
- Gagliardi D., Stepien P. P., Temperley R. J., Lightowlers R. N. und Chrzanowska-Lightowlers Z. M. A. 2004. Messenger RNA stability in mitochondria: different means to an end. *Trends in Genetics* 20: 260-67, <https://doi.org/10.1016/j.tig.2004.04.006>
- García-Mauriño S. M., Díaz-Quintana A., Rivero-Rodríguez F., Cruz-Gallardo I., Grüttner C., Hernández-Vellisca M. und Díaz-Moreno I. 2017. A putative RNA binding protein from *Plasmodium vivax* apicoplast. *FEBS Open Bio* 8: 177-88, <https://doi.org/10.1002/2211-5463.12351>
- Gardner M. J., Feagin J. E., Moore D. J., Spencer D. F., W.Gray M., Williamson D. H. und J.M.Wilson R. 1991. Organisation and expression of small subunit ribosomal RNA genes encoded by a 35-kilobase circular DNA in *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 48: 77-88, [https://doi.org/10.1016/0166-6851\(91\)90166-4](https://doi.org/10.1016/0166-6851(91)90166-4)
- Gardner M. J., et al. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419: 498-511, <https://doi.org/10.1038/nature01097>
- Ghezzi D., Saada A., DAdamo P., Fernandez-Vizarra E., Gasparini P., Tiranti V., Elpeleg O. und Zeviani M. 2008. FASTKD2 nonsense mutation in an infantile mitochondrial encephalomyopathy associated with cytochrome c oxidase deficiency. *American journal of human genetics* 83: 415-23, <https://doi.org/10.1016/j.ajhg.2008.08.009>
- Glaser S., van Dooren G. G., Agrawal S., Brooks C. F., McFadden G. I., Striepen B. und Higgins M. K. 2012. Tic22 is an essential chaperone required for protein import into the apicoplast. *The Journal of biological chemistry* 287: 39505-12, <https://doi.org/10.1074/jbc.M112.405100>
- Glushakova S., Humphrey G., Leikina E., Balaban A., Miller J. und Zimmerberg J. 2010. New stages in the program of malaria parasite egress imaged in normal and sickle erythrocytes. *Current biology : CB* 20: 1117-21, <https://doi.org/10.1016/j.cub.2010.04.051>
- Goodman C. D., Buchanan H. D. und McFadden G. I. 2017. Is the Mitochondrion a Good Malaria Drug Target? *Trends in Parasitology* 33: 185-93, <https://doi.org/10.1016/j.pt.2016.10.002>
- Goodman C. D., Su V. und McFadden G. I. 2007. The effects of anti-bacterials on the malaria parasite *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 152: 181-91, <https://doi.org/10.1016/j.molbiopara.2007.01.005>
- Groves M. R., Hanlon N., Turowski P., Hemmings B. A. und Barford D. 1999. The Structure of the Protein Phosphatase 2A PR65/A Subunit Reveals the

- Conformation of Its 15 Tandemly Repeated HEAT Motifs. *Cell* 96: 99-110, [https://doi.org/10.1016/S0092-8674\(00\)80963-0](https://doi.org/10.1016/S0092-8674(00)80963-0)
- Gupta A., Shah P., Haider A., Gupta K., Siddiqi M. I., Ralph S. A. und Habib S. 2014. Reduced ribosomes of the apicoplast and mitochondrion of *Plasmodium spp.* and predicted interactions with antibiotics. *Open biology* 4: 140045-45, <https://doi.org/10.1098/rsob.140045>
- Haïli N., Planchard N., Arnal N., Quadrado M., Vrielynck N., Dahan J., Francs-Small C. C. d. und Mireau H. 2016. The MTL1 Pentatricopeptide Repeat Protein Is Required for Both Translation and Splicing of the Mitochondrial NADH DEHYDROGENASE SUBUNIT7 mRNA in *Arabidopsis*. *Plant Physiology* 170: 354, <https://doi.org/10.1104/pp.15.01591>
- Hall N., et al. 2005. A Comprehensive Survey of the *Plasmodium* Life Cycle by Genomic, Transcriptomic, and Proteomic Analyses. *Science* 307: 82, <https://doi.org/10.1126/science.1103717>
- Hammani K., Bonnard G., Bouchoucha A., Gobert A., Pinker F., Salinas T. und Giegé P. 2014. Helical repeats modular proteins are major players for organelle gene expression. *Biochimie* 100: 141-50, <https://doi.org/10.1016/j.biochi.2013.08.031>
- Hammani K., Takenaka M., Miranda R. und Barkan A. 2016. A PPR protein in the PLS subfamily stabilizes the 5-end of processed *rpl16* mRNAs in maize chloroplasts. *Nucleic Acids Research* 44: 4278-88, <https://doi.org/10.1093/nar/gkw270>
- Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* 166: 557-80, [https://doi.org/10.1016/S0022-2836\(83\)80284-8](https://doi.org/10.1016/S0022-2836(83)80284-8)
- Henriquez F. L., Richards T. A., Roberts F., McLeod R. und Roberts C. W. 2005. The unusual mitochondrial compartment of *Cryptosporidium parvum*. *Trends in Parasitology* 21: 68-74, <https://doi.org/10.1016/j.pt.2004.11.010>
- Hermann T. 2000. Strategies for the Design of Drugs Targeting RNA and RNA-Protein Complexes. *Angewandte Chemie International Edition* 39: 1890-904, [https://doi.org/10.1002/1521-3773\(20000602\)39:11<1890::AID-ANIE1890>3.0.CO;2-D](https://doi.org/10.1002/1521-3773(20000602)39:11<1890::AID-ANIE1890>3.0.CO;2-D)
- Hicks J., Lassadi I., Carpenter E., Eno M., Vardakis A., Waller R. F., Howe C. und Nisbet R. E. R. 2018. A Pentatricopeptide Repeat Protein in the *Plasmodium* apicoplast is essential and shows sequence-specific RNA binding. *bioRxiv* <https://doi.org/10.1101/388728>
- Hillebrand A., Matz J. M., Almendinger M., Müller K., Matuschewski K. und Schmitz-Linneweber C. 2018. Identification of clustered organellar short (cos) RNAs and of a conserved family of organellar RNA-binding proteins, the

- heptatricopeptide repeat proteins, in the malaria parasite. *Nucleic Acids Research* 46: 10417-31, <https://doi.org/10.1093/nar/gky710>
- Hopkins J., Fowler R., Krishna S., Wilson I., Mitchell G. und Bannister L. 1999. The Plastid in *Plasmodium falciparum* Asexual Blood Stages: a Three-Dimensional Ultrastructural Analysis. *Protist* 150: 283-95, [https://doi.org/10.1016/S1434-4610\(99\)70030-1](https://doi.org/10.1016/S1434-4610(99)70030-1)
- Huber A. H., Nelson W. J. und Weis W. I. 1997. Three-Dimensional Structure of the Armadillo Repeat Region of  $\beta$ -Catenin. *Cell* 90: 871-82, [https://doi.org/10.1016/S0092-8674\(00\)80352-9](https://doi.org/10.1016/S0092-8674(00)80352-9)
- Ibba M. und Söll D. 2000. Aminoacyl-tRNA Synthesis. *Annual Review of Biochemistry* 69: 617-50, <https://doi.org/10.1146/annurev.biochem.69.1.617>
- Imanian B. und Keeling P. J. 2007. The dinoflagellates *Durinskia baltica* and *Kryptoperidinium foliaceum* retain functionally overlapping mitochondria from two evolutionarily distinct lineages. *BMC Evolutionary Biology* 7: 172, <https://doi.org/10.1186/1471-2148-7-172>
- Imwong M., Suwannasin K., Kunasol C., Sutawong K., Mayxay M., Rekol H., Smithuis F. M., Hlaing T. M., Tun K. M., van der Pluijm R. W., Tripura R., Miotto O., Menard D., Dhorda M., Day N. P. J., White N. J. und Dondorp A. M. 2017. The spread of artemisinin-resistant *Plasmodium falciparum* in the Greater Mekong subregion: a molecular epidemiology observational study. *The Lancet Infectious Diseases* 17: 491-97, [https://doi.org/10.1016/S1473-3099\(17\)30048-8](https://doi.org/10.1016/S1473-3099(17)30048-8)
- Jackson C. J., Norman J. E., Schnare M. N., Gray M. W., Keeling P. J. und Waller R. F. 2007. Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. *BMC Biology* 5: 41, <https://doi.org/10.1186/1741-7007-5-41>
- Janse C. J., Franke-Fayard B., Mair G. R., Ramesar J., Thiel C., Engelmann S., Matuschewski K., Gemert G. J. v., Sauerwein R. W. und Waters A. P. 2006. High efficiency transfection of *Plasmodium berghei* facilitates novel selection procedures. *Molecular and Biochemical Parasitology* 145: 60-70, <https://doi.org/10.1016/j.molbiopara.2005.09.007>
- Ji Y.-e., Mericle B. L., Rehkopf D. H., Anderson J. D. und Feagin J. E. 1996. The *Plasmodium falciparum* 6 kb element is polycistronically transcribed. *Molecular and Biochemical Parasitology* 81: 211-23, [http://dx.doi.org/10.1016/0166-6851\(96\)02712-0](http://dx.doi.org/10.1016/0166-6851(96)02712-0)
- Jourdain Alexis A., Koppen M., Rodley Christopher D., Maundrell K., Gueguen N., Reynier P., Guaras Adela M., Enriquez José A., Anderson P., Simarro M. und Martinou J.-C. 2015. A Mitochondria-Specific Isoform of FASTK Is Present In Mitochondrial RNA Granules and Regulates Gene Expression and Function. *Cell Reports* 10: 1110-21, <https://doi.org/10.1016/j.celrep.2015.01.063>

- Jourdain A. A., Popow J., de la Fuente M. A., Martinou J.-C., Anderson P. und Simarro M. 2017. The FASTK family of proteins: emerging regulators of mitochondrial RNA biology. *Nucleic Acids Research* 45: 10941-47, <https://doi.org/10.1093/nar/gkx772>
- Kapust R. B. und Waugh D. S. 1999. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* 8: 1668-74, <https://doi.org/10.1110/ps.8.8.1668>
- Ke H., Dass S., Morrissey J. M., Mather M. W. und Vaidya A. B. 2018. The mitochondrial ribosomal protein L13 is critical for the structural and functional integrity of the mitochondrion in *Plasmodium falciparum*. *Journal of Biological Chemistry* <https://doi.org/10.1074/jbc.RA118.002552>
- Ke H., Lewis Ian A., Morrissey Joanne M., McLean Kyle J., Ganesan Suresh M., Painter Heather J., Mather Michael W., Jacobs-Lorena M., Llinás M. und Vaidya Akhil B. 2015. Genetic Investigation of Tricarboxylic Acid Metabolism during the *Plasmodium falciparum* Life Cycle. *Cell Reports* 11: 164-74, <https://doi.org/10.1016/j.celrep.2015.03.011>
- Ke H., Morrissey J. M., Ganesan S. M., Mather M. W. und Vaidya A. B. 2012. Mitochondrial RNA polymerase is an essential enzyme in erythrocytic stages of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 185: 48-51, <https://doi.org/10.1016/j.molbiopara.2012.05.001>
- Ke H., Sigala P. A., Miura K., Morrissey J. M., Mather M. W., Crowley J. R., Henderson J. P., Goldberg D. E., Long C. A. und Vaidya A. B. 2014. The Heme Biosynthesis Pathway Is Essential for *Plasmodium falciparum* Development in Mosquito Stage but Not in Blood Stages. *Journal of Biological Chemistry* 289: 34827-37, <https://doi.org/10.1074/jbc.M114.615831>
- Klein D., Kern R. M. und Sokol R. Z. 1995. A method for quantification and correction of proteins after transfer to immobilization membranes. *Biochemistry and molecular biology international* 36: 59-66, PMID:7545052
- Kleinknecht L., Wang F., Stübe R., Philippar K., Nickelsen J. und Böhne A.-V. 2014. RAP, the Sole Octatricopeptide Repeat Protein in *Arabidopsis* Is Required for Chloroplast 16S rRNA Maturation. *The Plant Cell* 26: 777, <https://doi.org/10.1105/tpc.114.122853>
- Kleinknecht L., Wang F., Stübe R., Philippar K., Nickelsen J. und Böhne A.-V. 2016. CORRECTION: RAP, the Sole Octatricopeptide Repeat Protein in *Arabidopsis* Is Required for Chloroplast 16S rRNA Maturation. *The Plant Cell* 28: 984, <https://doi.org/10.1105/tpc.16.00094>
- Kobayashi T., Sato S., Takamiya S., Komaki-Yasuda K., Yano K., Hirata A., Onitsuka I., Hata M., Mi-ichi F., Tanaka T., Hase T., Miyajima A., Kawazu S.-i., Watanabe Y.-i. und Kita K. 2007. Mitochondria and apicoplast of *Plasmodium falciparum*: Behaviour on subcellular fractionation and the implication. *Mitochondrion* 7: 125-32, <http://dx.doi.org/10.1016/j.mito.2006.11.021>



- Kooij T. W. A. und Matuschewski K. 2007. Triggers and tricks of *Plasmodium* sexual development. *Current Opinion in Microbiology* 10: 547-53, <https://doi.org/10.1016/j.mib.2007.09.015>
- Kooij T. W. A., Rauch M. M. und Matuschewski K. 2012. Expansion of experimental genetics approaches for *Plasmodium berghei* with versatile transfection vectors. *Molecular and Biochemical Parasitology* 185: 19-26, <http://dx.doi.org/10.1016/j.molbiopara.2012.06.001>
- Korsinczky M., Chen N., Kotecka B., Saul A., Rieckmann K. und Cheng Q. 2000. Mutations in *Plasmodium falciparum* cytochrome b that are associated with atovaquone resistance are located at a putative drug-binding site. *Antimicrobial Agents and Chemotherapy* 44: 2100, <https://doi.org/10.1128/aac.44.8.2100-2108.2000>
- Kotera E., Tasaka M. und Shikanai T. 2005. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* 433: 326, <https://doi.org/10.1038/nature03229>
- Krungkrai S. R. und Krungkrai J. 2016. Insights into the pyrimidine biosynthetic pathway of human malaria parasite *Plasmodium falciparum* as chemotherapeutic target. *Asian Pacific Journal of Tropical Medicine* 9: 525-34, <https://doi.org/10.1016/j.apjtm.2016.04.012>
- Laemmli U. K. 1970. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 227: 680, <https://doi.org/10.1038/227680a0>
- Langmead B., Trapnell C., Pop M. und Salzberg S. L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10: R25-R25, <https://doi.org/10.1186/gb-2009-10-3-r25>
- Lapointe C. P., Wilinski D., Saunders H. A. J. und Wickens M. 2015. Protein-RNA networks revealed through covalent RNA marks. *Nature Methods* 12: 1163, <https://doi.org/10.1038/nmeth.3651>
- Lechner M., Rossmanith W., Hartmann R. K., Thölken C., Gutmann B., Giegé P. und Gobert A. 2015. Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. *Molecular Biology and Evolution* 32: 3186-93, <https://doi.org/10.1093/molbev/msv187>
- Lee A. H., Symington L. S. und Fidock D. A. 2014. DNA repair mechanisms and their biological roles in the malaria parasite *Plasmodium falciparum*. *Microbiology and Molecular Biology Reviews* 78: 469, <https://doi.org/10.1128/MMBR.00059-13>
- Lee I. und Hong W. 2004. RAP – a putative RNA-binding domain. *Trends in Biochemical Sciences* 29: 567-70, <https://doi.org/10.1016/j.tibs.2004.09.005>

- Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. und Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078-79, <https://doi.org/10.1093/bioinformatics/btp352>
- Li J., Maga J. A., Cermakian N., Cedergren R. und Feagin J. E. 2001. Identification and characterization of a *Plasmodium falciparum* RNA polymerase gene with similarity to mitochondrial RNA polymerases. *Molecular and Biochemical Parasitology* 113: 261-69, [https://doi.org/10.1016/S0166-6851\(01\)00223-7](https://doi.org/10.1016/S0166-6851(01)00223-7)
- Liere K., Weihe A. und Börner T. 2011. The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. *Journal of Plant Physiology* 168: 1345-60, <https://doi.org/10.1016/j.jplph.2011.01.005>
- Lightowlers R. N. und Chrzanowska-Lightowlers Z. M. A. 2013. Human pentatricopeptide proteins: only a few and what do they do? *RNA Biology* 10: 1433-38, <https://doi.org/10.4161/rna.24770>
- Lim L. und McFadden G. I. 2010. The evolution, metabolism and functions of the apicoplast. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 365: 749-63, <https://doi.org/10.1098/rstb.2009.0273>
- Lurin C., Andrés C., Aubourg S., Bellaoui M., Bitton F., Bruyère C., Caboche M., Debast C., Gualberto J., Hoffmann B., Lecharny A., Le Ret M., Martin-Magniette M.-L., Mireau H., Peeters N., Renou J.-P., Szurek B., Taconnat L. und Small I. 2004. Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *The Plant Cell* 16: 2089, <https://doi.org/10.1105/tpc.104.022236>
- Mak A. N.-S., Bradley P., Bogdanove A. J. und Stoddard B. L. 2013. TAL effectors: function, structure, engineering and applications. *Current opinion in structural biology* 23: 93-99, <https://doi.org/10.1016/j.sbi.2012.11.001>
- Manavski N., Guyon V., Meurer J., Wienand U. und Brettschneider R. 2012. An Essential Pentatricopeptide Repeat Protein Facilitates 5' Maturation and Translation Initiation of *rps3* mRNA in Maize Mitochondria. *The Plant Cell* 24: 3087, <https://doi.org/10.1105/tpc.112.099051>
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data Analysis* <https://doi.org/10.14806/ej.17.1.200>
- Mathews D. H., Sabina J., Zuker M. und Turner D. H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology* 288: 911-40, <https://doi.org/10.1006/jmbi.1999.2700>

- Matz J. M., Goosmann C., Matuschewski K. und Kooij T. W. A. 2018. An Unusual Prohibitin Regulates Malaria Parasite Mitochondrial Membrane Potential. *Cell Reports* 23: 756-67, <https://doi.org/10.1016/j.celrep.2018.03.088>
- McDermott J. J., Civic B. und Barkan A. 2018. Effects of RNA structure and salt concentration on the affinity and kinetics of interactions between pentatricopeptide repeat proteins and their RNA ligands. *PLOS ONE* 13: e0209713, <https://doi.org/10.1371/journal.pone.0209713>
- McFadden G. I., Reith M. E., Munholland J. und Lang-Unnasch N. 1996. Plastid in human parasites. *Nature* 381: 482, <https://doi.org/10.1038/381482a0>
- McMahon Aoife C., Rahman R., Jin H., Shen James L., Fieldsend A., Luo W. und Rosbash M. 2016. TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. *Cell* 165: 742-53, <https://doi.org/10.1016/j.cell.2016.03.007>
- Meierhoff K., Felder S., Nakamura T., Bechtold N. und Schuster G. 2003. HCF152, an *Arabidopsis* RNA Binding Pentatricopeptide Repeat Protein Involved in the Processing of Chloroplast *psbB-psbT-psbH-petB-petD* RNAs. *The Plant Cell* 15: 1480, <https://doi.org/10.1105/tpc.010397>
- Mingler M. K., Hingst A. M., Clement S. L., Yu L. E., Reifur L. und Koslowsky D. J. 2006. Identification of pentatricopeptide repeat proteins in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 150: 37-45, <https://doi.org/10.1016/j.molbiopara.2006.06.006>
- Mota M. M., Pradel G., Vanderberg J. P., Hafalla J. C. R., Frevert U., Nussenzweig R. S., Nussenzweig V. und Rodríguez A. 2001. Migration of *Plasmodium* Sporozoites Through Cells Before Infection. *Science* 291: 141, <https://doi.org/10.1126/science.291.5501.141>
- Mukherjee A. und Sadhukhan G. C. 2016. Anti-malarial Drug Design by Targeting Apicoplasts: New Perspectives. *Journal of pharmacopuncture* 19: 7-15, <https://doi.org/10.3831/KPI.2016.19.001>
- Murakami S., Kuehnle K. und Stern D. B. 2005. A spontaneous tRNA suppressor of a mutation in the *Chlamydomonas reinhardtii* nuclear MCD1 gene required for stability of the chloroplast *petD* mRNA. *Nucleic Acids Research* 33: 3372-80, <https://doi.org/10.1093/nar/gki651>
- Muralidharan V. und Goldberg D. E. 2013. Asparagine repeats in *Plasmodium falciparum* proteins: good for nothing? *PLOS Pathogens* 9: e1003488-e88, <https://doi.org/10.1371/journal.ppat.1003488>
- Nash E. A., Barbrook A. C., Howe C. J., Bernhardt K., Nisbet R. E. R. und Edwards-Stuart R. K. 2007. Organization of the Mitochondrial Genome in the Dinoflagellate *Amphidinium carterae*. *Molecular Biology and Evolution* 24: 1528-36, <https://doi.org/10.1093/molbev/msm074>

- Nisbet R. E. R., Kurniawan D. P., Bowers H. D. und Howe C. J. 2016. Transcripts in the *Plasmodium* Apicoplast Undergo Cleavage at tRNAs and Editing, and Include Antisense Sequences. *Protist* 167: 377-88, <https://doi.org/10.1016/j.protis.2016.06.003>
- Nisbet R. E. R. und McKenzie J. L. 2016. Transcription of the apicoplast genome. *Molecular and Biochemical Parasitology* 210: 5-9, <https://doi.org/10.1016/j.molbiopara.2016.07.004>
- Nishikura K. 2015. A-to-I editing of coding and non-coding RNAs by ADARs. *Nature Reviews Molecular Cell Biology* 17: 83, <https://doi.org/10.1038/nrm.2015.4>
- Nixon G. L., Pidathala C., Shone A. E., Antoine T., Fisher N., O'Neill P. M., Ward S. A. und Biagini G. A. 2013. Targeting the mitochondrial electron transport chain of *Plasmodium falciparum*: new strategies towards the development of improved antimalarials for the elimination era. *Future Medicinal Chemistry* 5: 1573-91, <https://doi.org/10.4155/fmc.13.121>
- Nordberg H., Cantor M., Dusheyko S., Hua S., Poliakov A., Shabalov I., Smirnova T., Grigoriev I. V. und Dubchak I. 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Research* 42: D26-D31, <https://doi.org/10.1093/nar/gkt1069>
- Nosten F. und White N. J. 2007. Artemisinin-Based Combination Treatment of Falciparum Malaria. *The American journal of tropical medicine and hygiene* 77: 181-92, <https://doi.org/10.4269/ajtmh.2007.77.181>
- Ojala D., Montoya J. und Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290: 470, <https://doi.org/10.1038/290470a0>
- Painter H. J., Morrissey J. M., Mather M. W. und Vaidya A. B. 2007. Specific role of mitochondrial electron transport in blood-stage *Plasmodium falciparum*. *Nature* 446: 88-91, <https://doi.org/10.1038/nature05572>
- Pall G. S., Codony-Servat C., Byrne J., Ritchie L. und Hamilton A. 2007. Carbodiimide-mediated cross-linking of RNA to nylon membranes improves the detection of siRNA, miRNA and piRNA by northern blot. *Nucleic Acids Research* 35: e60-e60, <https://doi.org/10.1093/nar/gkm112>
- Pall G. S. und Hamilton A. J. 2008. Improved northern blot method for enhanced detection of small RNA. *Nature Protocols* 3: 1077, <https://doi.org/10.1038/nprot.2008.67>
- Payne D. 1987. Spread of chloroquine resistance in *Plasmodium falciparum*. *Parasitology Today* 3: 241-46, [https://doi.org/10.1016/0169-4758\(87\)90147-5](https://doi.org/10.1016/0169-4758(87)90147-5)

- Petersen I., Eastman R. und Lanzer M. 2011. Drug-resistant malaria: Molecular mechanisms and implications for public health. *FEBS Letters* 585: 1551-62, <https://doi.org/10.1016/j.febslet.2011.04.042>
- Pfalz J., Bayraktar O. A., Prikryl J. und Barkan A. 2009. Site-specific binding of a PPR protein defines and stabilizes 5 and 3 mRNA termini in chloroplasts. *The EMBO journal* 28: 2042-52, <https://doi.org/10.1038/emboj.2009.121>
- Prikryl J., Rojas M., Schuster G. und Barkan A. 2011. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proceedings of the National Academy of Sciences of the United States of America* 108: 415-20, <https://doi.org/10.1073/pnas.1012076108>
- Prudêncio M., Rodriguez A. und Mota M. M. 2006. The silent path to thousands of merozoites: the *Plasmodium* liver stage. *Nature Reviews Microbiology* 4: 849, <https://doi.org/10.1038/nrmicro1529>
- Quinlan A. R. und Hall I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841-42, <https://doi.org/10.1093/bioinformatics/btq033>
- Raabe C. A., Sanchez C. P., Randau G., Robeck T., Skryabin B. V., Chinni S. V., Kube M., Reinhardt R., Ng G. H., Manickam R., Kuryshev V. Y., Lanzer M., Brosius J., Tang T. H. und Rozhdestvensky T. S. 2010. A global view of the nonprotein-coding transcriptome in *Plasmodium falciparum*. *Nucleic Acids Research* 38: 608-17, <https://doi.org/10.1093/nar/gkp895>
- Rahire M., Laroche F., Cerutti L. und Rochaix J.-D. 2012. Identification of an OPR protein involved in the translation initiation of the *PsaB* subunit of photosystem I. *The Plant Journal* 72: 652-61, <https://doi.org/10.1111/j.1365-313X.2012.05111.x>
- Ralph S. A., D'Ombra M. C. und McFadden G. I. 2001. The apicoplast as an antimalarial drug target. *Drug Resistance Updates* 4: 145-51, <https://doi.org/10.1054/drup.2001.0205>
- Ralph S. A., van Dooren G. G., Waller R. F., Crawford M. J., Fraunholz M. J., Foth B. J., Tonkin C. J., Roos D. S. und McFadden G. I. 2004. Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nature Reviews Microbiology* 2: 203, <https://doi.org/10.1038/nrmicro843>
- Rehkopf D. H., Gillespie D. E., Harrell M. I. und Feagin J. E. 2000. Transcriptional mapping and RNA processing of the *Plasmodium falciparum* mitochondrial mRNAs. *Molecular and Biochemical Parasitology* 105: 91-103, [https://doi.org/10.1016/S0166-6851\(99\)00170-X](https://doi.org/10.1016/S0166-6851(99)00170-X)
- Roy A., Kucukural A. und Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* 5: 725, <https://doi.org/10.1038/nprot.2010.5>

- Ruwe H., Gutmann B., Schmitz-Linneweber C., Small I. und Kindgren P. 2018. The E-domain of CRR2 participates in sequence-specific recognition of RNA in plastids. *New Phytologist* 0, <https://doi.org/10.1111/nph.15578>
- Ruwe H. und Schmitz-Linneweber C. 2012. Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? *Nucleic Acids Research* 40: 3106-16, <https://doi.org/10.1093/nar/gkr1138>
- Ruwe H., Wang G., Gusewski S. und Schmitz-Linneweber C. 2016. Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. *Nucleic Acids Research* 44: 7406-17, <https://doi.org/10.1093/nar/gkw466>
- Sambrook J. und Russell D. W. 2001. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press* 2001 <https://doi.org/10.1101/pdb.prot4020>
- Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden C., Saalfeld S., Schmid B., Tinevez J.-Y., White D. J., Hartenstein V., Eliceiri K., Tomancak P. und Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Meth* 9: 676-82, <https://doi.org/10.1038/nmeth.2019>
- Schmitz-Linneweber C. und Small I. 2008. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends in Plant Science* 13: 663-70, <https://doi.org/10.1016/j.tplants.2008.10.001>
- Schwach F., Bushell E., Gomes A. R., Anar B., Girling G., Herd C., Rayner J. C. und Billker O. 2015. PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites. *Nucleic Acids Research* 43: D1176-D82, <https://doi.org/10.1093/nar/gku1143>
- Schwarz C., Elles I., Kortmann J., Piotrowski M. und Nickelsen J. 2007. Synthesis of the D2 Protein of Photosystem II in *Chlamydomonas* Is Controlled by a High Molecular Mass Complex Containing the RNA Stabilization Factor Nac2 and the Translational Activator RBP40. *The Plant Cell* 19: 3627, <https://doi.org/10.1105/tpc.107.051722>
- Sharma A. und Sharma A. 2015. *Plasmodium falciparum* mitochondria import tRNAs along with an active phenylalanyl-tRNA synthetase. *Biochemical Journal* 465: 459, <https://doi.org/10.1042/BJ20140998>
- Shoguchi E., Shinzato C., Hisata K., Satoh N. und Mungpakdee S. 2015. The Large Mitochondrial Genome of *Symbiodinium minutum* Reveals Conserved Noncoding Sequences between Dinoflagellates and Apicomplexans. *Genome biology and evolution* 7: 2237-44, <https://doi.org/10.1093/gbe/evv137>
- Sidik S. M., Huet D., Ganesan S. M., Huynh M.-H., Wang T., Nasamu A. S., Thiru P., Saeij J. P. J., Carruthers V. B., Niles J. C. und Lourido S. 2016. A Genome-

- wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes. *Cell* 166: 1423-35.e12, <https://doi.org/10.1016/j.cell.2016.08.019>
- Simarro M., Gimenez-Cassina A., Kedersha N., Lazaro J.-B., Adelmant G. O., Marto J. A., Rhee K., Tisdale S., Danial N., Benarafa C., Orduña A. und Anderson P. 2010. Fast kinase domain-containing protein 3 is a mitochondrial protein essential for cellular respiration. *Biochemical and Biophysical Research Communications* 401: 440-46, <https://doi.org/10.1016/j.bbrc.2010.09.075>
- Sinnis P. und Zavala F. 2012. The skin: where malaria infection and the host immune response begin. *Seminars in Immunopathology* 34: 787-92, <https://doi.org/10.1007/s00281-012-0345-5>
- Slamovits C. 2014. Mitochondrial Genomes in Alveolates. *Molecular Life Sciences: An Encyclopedic Reference*: 1-7, [https://doi.org/10.1007/978-1-4614-6436-5\\_117-3](https://doi.org/10.1007/978-1-4614-6436-5_117-3)
- Small I. D. und Peeters N. 2000. The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends in Biochemical Sciences* 25: 45-47, [http://dx.doi.org/10.1016/S0968-0004\(99\)01520-0](http://dx.doi.org/10.1016/S0968-0004(99)01520-0)
- Soll J. und Schleiff E. 2004. Protein import into chloroplasts. *Nature Reviews Molecular Cell Biology* 5: 198, <https://doi.org/10.1038/nrm1333>
- Spork S., Hiss J. A., Mandel K., Sommer M., Kooij T. W. A., Chu T., Schneider G., Maier U. G. und Przyborski J. M. 2009. An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. *Eukaryotic cell* 8: 1134-45, <https://doi.org/10.1128/EC.00083-09>
- Srivastava I. K. und Vaidya A. B. 1999. A Mechanism for the Synergistic Antimalarial Action of Atovaquone and Proguanil. *Antimicrobial Agents and Chemotherapy* 43: 1334, <https://doi.org/10.1128/AAC.43.6.1334>
- Sturm A., Mollard V., Cozijnsen A., Goodman C. D. und McFadden G. I. 2015. Mitochondrial ATP synthase is dispensable in blood-stage *Plasmodium berghei* rodent malaria but essential in the mosquito phase. *Proceedings of the National Academy of Sciences* 112: 10216, <https://doi.org/10.1073/pnas.1423959112>
- Su X.-z., Wu Y., Sifri C. D. und Wellems T. E. 1996. Reduced Extension Temperatures Required for PCR Amplification of Extremely A+T-rich DNA. *Nucleic Acids Research* 24: 1574-75, <https://doi.org/10.1093/nar/24.8.1574>
- Sugita M. und Ichinose M. 2018. The DYW Domains of Pentatricopeptide Repeat RNA Editing Factors Contribute to Discriminate Target and Non-Target Editing Sites. *Plant and Cell Physiology* 59: 1652-59, <https://doi.org/10.1093/pcp/pcy086>
- Taanman J.-W. 1999. The mitochondrial genome: structure, transcription, translation and replication. *Biochimica et Biophysica Acta (BBA)* -

- Bioenergetics* 1410: 103-23, [https://doi.org/10.1016/S0005-2728\(98\)00161-3](https://doi.org/10.1016/S0005-2728(98)00161-3)
- Takashima E., Takamiya S., Takeo S., Mi-ichi F., Amino H. und Kita K. 2001. Isolation of mitochondria from *Plasmodium falciparum* showing dihydroorotate dependent respiration. *Parasitology International* 50: 273-78, [http://dx.doi.org/10.1016/S1383-5769\(01\)00085-X](http://dx.doi.org/10.1016/S1383-5769(01)00085-X)
- Tian Q., Taupin J., Elledge S., Robertson M. und Anderson P. 1995. Fas-activated serine/threonine kinase (FAST) phosphorylates TIA-1 during Fas-mediated apoptosis. *The Journal of Experimental Medicine* 182: 865, <https://doi.org/10.1084/jem.182.3.865>
- Trager W. und Jensen J. B. 1976. Human malaria parasites in continuous culture. *Science* 193: 673, <https://doi.org/10.1126/science.781840>
- Vaidya A. B. und Arasu P. 1987. Tandemly arranged gene clusters of malarial parasites that are highly conserved and transcribed. *Molecular and Biochemical Parasitology* 22: 249-57, [https://doi.org/10.1016/0166-6851\(87\)90056-9](https://doi.org/10.1016/0166-6851(87)90056-9)
- van Dooren G. G., Marti M., Tonkin C. J., Stimmler L. M., Cowman A. F. und McFadden G. I. 2005. Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of *Plasmodium falciparum*. *Molecular Microbiology* 57: 405-19, <https://doi.org/10.1111/j.1365-2958.2005.04699.x>
- van Dooren G. G., Stimmler L. M. und McFadden G. I. 2006. Metabolic maps and functions of the *Plasmodium* mitochondrion. *FEMS Microbiology Reviews* 30: 596-630, <https://doi.org/10.1111/j.1574-6976.2006.00027.x>
- Waller R. F. und Jackson C. J. 2009. Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. *BioEssays* 31: 237-45, <https://doi.org/10.1002/bies.200800164>
- Walter A. E., Turner D. H., Kim J., Lyttle M. H., Müller P., Mathews D. H. und Zuker M. 1994. Coaxial stacking of helices enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proceedings of the National Academy of Sciences* 91: 9218, <https://doi.org/10.1073/pnas.91.20.9218>
- Waltz F., Nguyen T.-T., Arrivé M., Bochler A., Chicher J., Hammann P., Kuhn L., Quadrado M., Mireau H., Hashem Y. und Giegé P. 2019. Small is big in *Arabidopsis* mitochondrial ribosome. *Nature Plants* 5: 106-17, <https://doi.org/10.1038/s41477-018-0339-y>
- Wang C., Aubé F., Planchard N., Quadrado M., Dargel-Graffin C., Nogué F. und Mireau H. 2017. The pentatricopeptide repeat protein MTSF2 stabilizes a *nad1* precursor transcript and defines the 3' end of its 5'-half intron. *Nucleic Acids Research* 45: 6119-34, <https://doi.org/10.1093/nar/gkx162>



- Wang F., Johnson X., Cavauiolo M., Bohne A.-V., Nickelsen J. und Vallon O. 2015. Two *Chlamydomonas* OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b6f. *The Plant Journal* 82: 861-73, <https://doi.org/10.1111/tpj.12858>
- Wang M., Ogé L., Perez-Garcia M.-D., Hamama L. und Sakr S. 2018. The PUF Protein Family: Overview on PUF RNA Targets, Biological Functions, and Post Transcriptional Regulation. *International journal of molecular sciences* 19: 410, <https://doi.org/10.3390/ijms19020410>
- Wellems T. E. und Plowe C. V. 2001. Chloroquine-Resistant Malaria. *The Journal of Infectious Diseases* 184: 770-76, <https://doi.org/10.1086/322858>
- WHO. 2017. World Malaria Report 2017, ISBN: 978 92 4 156552 3
- Wilson R. J. M., Denny P. W., Preiser P. R., Rangachari K., Roberts K., Roy A., Whyte A., Strath M., Moore D. J., Moore P. W. und Williamson D. H. 1996. Complete Gene Map of the Plastid-like DNA of the Malaria Parasite *Plasmodium falciparum*. *Journal of Molecular Biology* 261: 155-72, <https://doi.org/10.1006/jmbi.1996.0449>
- Wolf A. R. und Mootha V. K. 2014. Functional genomic analysis of human mitochondrial RNA processing. *Cell Reports* 7: 918-31, <https://doi.org/10.1016/j.celrep.2014.03.035>
- Wu W., Liu S., Ruwe H., Zhang D., Melonek J., Zhu Y., Hu X., Gusewski S., Yin P., Small I. D., Howell K. A. und Huang J. 2016. SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S–4.5S rRNA precursors in *Arabidopsis thaliana*. *The Plant Journal* 85: 607-21, <https://doi.org/10.1111/tpj.13126>
- Yeh E. und DeRisi J. L. 2011. Chemical Rescue of Malaria Parasites Lacking an Apicoplast Defines Organelle Function in Blood-Stage *Plasmodium falciparum*. *PLOS Biology* 9: e1001138, <https://doi.org/10.1371/journal.pbio.1001138>
- Zhang M., Wang C., Otto T. D., Oberstaller J., Liao X., Adapa S. R., Udenze K., Bronner I. F., Casandra D., Mayho M., Brown J., Li S., Swanson J., Rayner J. C., Jiang R. H. Y. und Adams J. H. 2018. Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science* 360, <https://doi.org/10.1126/science.aap7847>
- Zhelyazkova P., Hammani K., Rojas M., Voelker R., Vargas-Suárez M., Börner T. und Barkan A. 2012. Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. *Nucleic Acids Research* 40: 3092-105, <https://doi.org/10.1093/nar/gkr1137>
- Zhu G., Marchewka M. J. und Keithly J. S. 2000. *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology* 146: 315-21, <https://doi.org/10.1099/00221287-146-2-315>

- Zoschke R., Qu Y., Zubo Y. O., Börner T. und Schmitz-Linneweber C. 2013. Mutation of the pentatricopeptide repeat-SMR protein SVR7 impairs accumulation and translation of chloroplast ATP synthase subunits in *Arabidopsis thaliana*. *Journal of Plant Research* 126: 403-14, <https://doi.org/10.1007/s10265-012-0527-1>
- Zuegge J., Ralph S., Schmuker M., McFadden G. I. und Schneider G. 2001. Deciphering apicoplast targeting signals – feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. *Gene* 280: 19-26, [https://doi.org/10.1016/S0378-1119\(01\)00776-4](https://doi.org/10.1016/S0378-1119(01)00776-4)

## Anhang

### Hersteller und Anbieter

Anhang Tabelle 1: Liste der Anbieter von Chemikalien und Enzymen.

Firma	Sitz
Agilent	Agilent Technologies, Santa Clara, CA, USA
Baxter	Baxter Deutschland GmbH, Unterschleißheim, DE
Bio-Rad	Bio-Rad Laboratories, Hercules, CA, USA
Biozym	Biozym Scientific GmbH, Hessisch Oldendorf, DE
Carl Roth	Carl Roth GmbH & Co, Karlsruhe, DE
GE Healthcare	GE Healthcare Europe GmbH, Freiburg, DE
Hartmann Analytics	Hartmann Analytics GmbH, Braunschweig, DE
Illumina	Illumina Inc., San Diego, CA, USA
Life Technologies	Life Technologies, Calrsbad, CA, USA
Lonza	Lonza Group AG, Basel, CH
Merck	Merck KGaA, Darmstadt, DE
MP Biochemicals	MP Biochemicals, Santa Ana, CA, USA
Mytron	mytron Bio- und Solartechnik GmbH, Heilbad Heiligenstadt, DE
NEB	New England Biolabs, Ipswich, MA, USA
PAN-Biotech	PAN-Biotech GmbH, Aidenbach, DE
PEQLAB	PEQLAB, Erlangen, DE
Promega	Promega Corporation, Fitchburg, WI, USA
Qiagen	Qiagen, Holden, DE
Ratiopharm	Ratiopharm GmbH, Ulm, DE
Roboklon	Robokol GmbH, Berlin, DE
Roche	Roche Diagnostics GmbH, Mannheim, DE
Sigma-Aldrich	Sigma-Aldrich Corporation, St. Louis, MO, USA
SMB	Services in Molecular Biology GmbH, Rüdersdorf, DE
Thermo Fisher	Thermo Fisher Scientific Inc., Waltham, MA, USA
Veolia	Veolia Water Solutions & Technologies, Saint Maurice, FR
Zymo Research	Zymo Research Corporation, Irvine, CA, USA

## Liste der verwendeten Oligonukleotide

Anhang Tabelle 2: Liste der verwendeten Oligonukleotide und deren Anwendung. Sequenzen sind in 5'-3'-Richtung angegeben. Oligonukleotide für die *in vitro* Transkription sind mit T7-Promotorsequenzen (unterstrichen) versehen. Anwendungen: qRT-PCR = Anreicherungsbestimmung in der Organellenpräparation; Northern-Sonde = für die radioaktive Detektion von cosRNAs; Klonierung = für die Amplifikation genomischer Bereiche für die Erstellung von verschiedenen Konstrukten; Kolonie-PCR = Insertionsprüfung für die Klonierung; Genotypisierung = Überprüfung des Integrationserfolges; i.v.Transkription = Erstellung von DNA-Vorlagen für die *in vitro* Transkription.

Name	Sequenz 5' nach 3'	Anwendung
<b>Organellenanreicherung und cosRNA-Detektion</b>		
Pf_A_sufB_fwd	TCGAACCTACTTCAACTTGTATCC	qRT-PCR
Pf_A_sufB_rev	AGGAGATTATTTAGGTAATGGTGGT	qRT-PCR
Pf_1_18S_fwd	ATGCTTTTATTGCTTTTGAGAGGT	qRT-PCR
Pf_1_18S_rev	ACTCAATCATGACTACCCGTCTG	qRT-PCR
Pf_mito.rRNA.SSUF	GTTACGACTTCGCACCGACTGTTT	Northern Sonde
Pfm_rRNA11 5	CAAATATTTCTCATTTATATCCG	Northern Sonde
<b>Markierungskonstrukte</b>		
Pb_fwd 081630 + Eag1	CGATCGGCCGACATGTTTCATCCCAAGGTTTAG	Klonierung
Pb_rev 081630	ATCATTATTTTTTGGGGAGCAAATC	Klonierung
Pb_fwd 051680 + Eag1	CGATCGGCCGGGCGATGGATGTATGCAC	Klonierung
Pb_rev 051680	TTTTTGATCTTGATTGGCTGATTTTTTATC	Klonierung
Pb_fwd 040990 + Eag1	CGTACGGCCGGCTCGACTCAAATTTGAAACAC	Klonierung
Pb_rev 040990	CTTAATTGATCTACACTTTTCAATCTG	Klonierung
KO PB 040990 3 fwd	TAGCCTCGAGCCAATTGGTAACTTTTGAAATAAT ATTAAATG	Klonierung
KO PB 040990 3rev	TAGCAAGCTTCATGGAAGTTAGGATATGCTGTTG	Klonierung
Pb_fwd 143670 + Eag1	CGTACGGCCGCTAGACACACAACATTTTTCTCAG	Klonierung
Pb_rev 143670	TGTATCTTTCAAAATATATTCATTGGTCC	Klonierung
KO PB 143670 3 rev	TAGCAAGCTTTTGCTAAGTCATATATATCCATACC TCAC	Klonierung
KO PB 143670 3 fwd	TAGCCTCGAGGCTTTTTTATTTGTTGAGCAATCT TG	Klonierung
Pb_fwd 051380 + Eag1	CGATCGGCCGGCGAAACGTATGCATAAGC	Klonierung
Pb_rev 051380	CATATTCAGTTCTATTCTTTCTTTG	Klonierung
KO PB 051380 3 fwd	TAGCCTCGAGATTTTTTAACTCATCTTAATTTAG TCTGAC	Klonierung
KO PB 051380 3 rev	TAGCAAGCTTAGATCCCCAATTATTTAATGACATT AACC	Klonierung
PB111800 5 fwd + Eag1	TCAGCGGCCGGATAAAGGAAAGGAGTTGTTATGT AG	Klonierung

## Anhang

PB111800 5 rev	AAATATTTTATTTTTTTTAACAACAATTTTTCCTG	Klonierung
PB111800 3 fwd + XhoI	TCAGCTCGAGTGGACACAATTAAACCACAGC	Klonierung
PB111800 3 rev + HindIII	TCAGAAAGCTTTTAGCTATTATGCAACCAAGCC	Klonierung
PB142250 5 fwd + EagI	TCAGCGGCCGATATAGCTCTAGTAGCAAATTCTTAGC	Klonierung
PB rev 142250 +HpaI	TACGGTTAACTTTTTCATGGTTCGGGGTATTATTC	Klonierung
KO PB 142250 3 fwd	TAGCCTCGAGATTTCTACATAGATCTACGCACCC	Klonierung
KO PB142250 3 rev ClaI	TAGCATCGATGGGCGTTATAGCTCGTTAAAC	Klonierung
pBAT1-G seq tag fwd	CGACGGCCAGTTAATACGACTCAC	Kolonie-PCR
pBAT1-G Insert seq rev	CGCATGAACTCCTTGATGATGGCC	Kolonie-PCR
pBAT1-G seq tag rev	AGCACAACAAAATGGATGATTCCCCTC	Kolonie-PCR
<b>Genotypisierung für Markierungskonstrukte</b>		
mCherry_rev_Geno	CCCTCCATGTGAACCTTGAAG	Genotypisierung
PB143670 3Geno In	ATCGGCTGAACAAACTGAAAC	Genotypisierung
PB143670 3Geno Out	GCTCAAACATCTAGATTTAAAGGTG	Genotypisierung
PB051680 3Geno In	CCTAAAAGGATAAAAAATCAGCCAATAC	Genotypisierung
PB051680 3Geno Out	CCCTCGAAATAAATTTCCCTTTTCC	Genotypisierung
PB142250 3Geno In	CCCGAACCATGAAAAATAATTTTATGG	Genotypisierung
PB142250 3Geno Out	GTGTGGTGGGAAAAAACATG	Genotypisierung
PB040990 3Geno In	CTAAGCAAGGATTTTCATCAATACCAG	Genotypisierung
PB040990 3Geno Out	CAACTACTCAATGACCCCCTTAC	Genotypisierung
PB051380 geno 3' g fwd	CCATTAGGAAAAAATTAATACACACTCC	Genotypisierung
PB051380 geno 3' rev	CTATTCTCATTTAGGTTGTCTATTTTGC	Genotypisierung
PB081630 3Geno In	TGTATGCAGATTGGATTTGCTCC	Genotypisierung
PB081630 3Geno Out	CATCTTTAAAGGATGAATGTGCAG	Genotypisierung
<b>Rekombinante Proteine und in vitro Transkription</b>		
pMAL-TEV PCR rev	GGATCCCTGGAAGTACAGG	Kolonie-PCR
pMAL-TEV fwd insert seq	CAACCTCGGGATCGAGGGAAG	Kolonie-PCR
pMAL-TEV rev insert seq	CAGGGTTTTCCAGTCACGACG	Kolonie-PCR
PB051680 rec sig- fwd	TATGAGGAAATAAAAAATTTAAAGGCACAAGAATTG	Klonierung
PB 051680 rec rev + SalI	TGACGTCGACCTATCCAATTTGGAAAGACAAAATTAGG	Klonierung
PB051380 rec sig- fwd	AATGTTTGCTTAACAAGTCAGGTAAATTTAATG	Klonierung
PB 051380 rec rev + Sall	TGACGTCGACGAGCAACTTAAAAATAAAGTTATTTTCGTG	Klonierung
PB_mito_1- T7	(TAATACGACTCACTATAGGG)TCAGGCGTTAAAA GCGTTC	i.v.Transkription
PB_mito_1+ T7	(TAATACGACTCACTATAGGG)CATAGAGACTCAG CCTATG	i.v.Transkription
PB_mito_2- T7	(TAATACGACTCACTATAGGG)CCATAAGAAATTAA TATTTAATGGACATGGG	i.v.Transkription
PB_mito_2+ T7	(TAATACGACTCACTATAGGG)TAACTACATTAAAG GAACTCGACTG	i.v.Transkription
PB_mito_3- T7	(TAATACGACTCACTATAGGG)AGAACTTCTTTAAC TTGTCAACTCCC	i.v.Transkription
PB_mito_3+ T7	(TAATACGACTCACTATAGGG)GTTCAATTGGATCA ACAATGACTTTATTG	i.v.Transkription
PB_mito_4 - T7	(TAATACGACTCACTATAGGG)AGGATTTAACGGT CAACCTATTATTGTTC	i.v.Transkription
PB_mito_4+ T7	(TAATACGACTCACTATAGGG)CGATATTATTACCG TACAAGCCG	i.v.Transkription

## Anhang

Mito probe 1 + str	CTGTCTAGATTAGACAATACATATTTGTGTATAAA GCTAGAGTAC	Northern Sonde
Mito Probe 1 -str	GTA CTCTAGCTTTATACACAAATATGTATTGTCTA ATCTAGACAG	Northern Sonde
Mito probe 2 + str	CGCCACCTGTAAAACTGGTAAAGTTAATAATAAC ATTATTGATG	Northern Sonde
Mito Probe 2 - str	CATCAATAATGTTATTATTAACCTTACCAGTTTTAA CAGGTGGCG	Northern Sonde
Mito probe 3 + str	GGGTATTTTAAATGCTGTATCATACCCTAAAGGAT TAGTGCTACC	Northern Sonde
Mito probe 3 -str	GGTAGCACTAATCCTTTAGGGTATGATACAGCATT AAAAATACCC	Northern Sonde
Mito probe 4 + str	GAATCCAATAGCGTTGATTAAGACATCGATATA CGGATTTCTC	Northern Sonde
Mito probe 4 - str	GAGAAATCCGTATATCGATGTCTTTAATCAACGC TATTGGATTG	Northern Sonde
<b>Deletionskonstrukte</b>		
KO PB 020810 3 rev	TAGCAAGCTTTGTTTGTATCAATAAGCCAGC	Klonierung
KO PB 020810 5 fwd	TAGCCAGCTGGTAACATGTATTTGATATATCGTTT CGTT	Klonierung
KO PB 020810 5 rev	TAGCTCTAGACATGGCTAGCAAACCAAGTAAG	Klonierung
KO PB 020810 3 fwd	TAGCCTCGAGGTAGAGCATCAAAATAAATAGGCA TG	Klonierung
KO PB 080730 3 fwd	TAGCCTCGAGCTTTTTTTGTATAGCTAGCTAAAA TTAAGG	Klonierung
KO PB 080730 3 rev	TCCTAAGCTTCCCATCCAAATG	Klonierung
KO PB 080730 5 fwd	TAGCTCTAGACATTTATTATTTCAGCTTTTCTCTCC G	Klonierung
KO PB 080730 5 rev	TAGCCAGCTGTATCCCTTGAGCCATTAAATGGC	Klonierung
KO PB 143670 3 rev	TAGCAAGCTTTTGCTAAGTCATATATATCCATACC TCAC	Klonierung
KO PB 143670 3 fwd	TAGCCTCGAGGCTTTTTTATTTGTTTCAGCAATCT TG	Klonierung
KO PB 143670 5 rev	TGACGTTAACGAAATGTGAACAAAAATACAGTTCC C	Klonierung
KO PB 143670 5 fwd	TAGCTCTAGACACACTATGGATTTGATGAAAATCA G	Klonierung
KO PB 051680 3 fwd	TAGCCTCGAGCACAAATAGTGCATGAAAAAATG AAAG	Klonierung
KO PB 051680 3 rev	TGACAAGCTTACACGAATAATACATTGTAAAGGAG G	Klonierung
KO PB 051680 5 fwd	TAGCTCTAGAATTAGTCCATTGCAATTTGCTAG	Klonierung
KO PB 051680 5 rev	TAGCCAGCTGCATTCTAATCAATTCAGTTAAATG GGG	Klonierung
KO PB 142250 3 fwd	TAGCCTCGAGATTTCTACATAGATCTACGCACCC	Klonierung
KO PB 142250 3 rev	TAGCAAGCTTGGGCGTTATAGCTCGTTAAAC	Klonierung
KO PB 142250 5 fwd	TAGCTCTAGAATGTTTCATTACCTTAACATAGTTG AATAGG	Klonierung
KO PB 142250 5 rev	TAGCCAGCTGTTTCATCCAAATTTGTACCTGC	Klonierung
KO PB 092630 3 fwd	TAGCCTCGAGACGTAATAAATGTTCCGCATGC	Klonierung
KO PB 092630 3 rev	TAGCATCGATATCTTTATTTCAGTAAAAAGCCCTC	Klonierung
KO PB 092630 5 fwd	TAGCTCTAGAGAGAAGAAGCTAATAGTATGATAA CTGC	Klonierung
KO PB 092630 5 rev	TAGCCAGCTGCCATTTGAATACGAAAACAAAATGT G	Klonierung
KO PB 124410 3 fwd	TAGCCTCGAGGCAATTGTGATTTTCGTATTAAATT CAATAG	Klonierung

## Anhang

KO PB 124410 3 rev	TAGCAAGCTTGTACGCACTTGAAATTATTACATAT TGAG	Klonierung
KO PB 124410 5 fwd	TAGCTCTAGATTCTCAATTTTCCAGGGGAGG	Klonierung
KO PB 124410 5 rev	TAGCCAGCTGAGCAGTTCAGAAAGTAAGTATGAA TAG	Klonierung
KO PB 081630 3 fwd	TAGCCTCGAGGATTTGGTGAATACCTTCAAAGA TG	Klonierung
KO PB 081630 3 rev	TAGCAAGCTTCAAGGGGATGCAGAGAATAATAG	Klonierung
KO PB 081630 5 fwd	TAGCTCTAGAGCATGCAGCTTCAATTGTTATAG	Klonierung
KO PB 081630 5 rev	TAGCCAGCTGAGCAAGCAAAGGAGCTATTTATC	Klonierung
KO PB 144130 3 fwd	TAGCCTCGAGGTAATCCATTGTGCTTATGCTACTA TG	Klonierung
KO PB 144130 3 rev	TAGCAAGCTTCAGAATTACCTGACAATTCACATTT G	Klonierung
KO PB 144130 5 fwd	TAGCTCTAGAGTTACTAAAAGGGCCTTAGTTAAGA C	Klonierung
KO PB 144130 5 rev	TAGCCAGCTGCCATGCATATACACCACGC	Klonierung
KO PB 040990 3 fwd	TAGCCTCGAGCCAATTGGTAACTTTGGAAATAAT ATTAAATG	Klonierung
KO PB 040990 3rev	TAGCAAGCTTCATGGAAGTTAGGATATGCTGTTG	Klonierung
KO PB 040990 5 fwd	TAGCTCTAGACGAATAGTCTAATTCAAACACAATC AC	Klonierung
KO PB 040990 5 rev	TAGCCAGCTGACAGTATCCTTATTTTGTCAATCGC	Klonierung
PB051380 KO 3' fwd	ATGCTCGAGAGAGAAACACGAAAAATAACTTTATT TTTAAAG	Klonierung
PB051380 KO 3' rev	ATGATCGATACATTAACCCAAGTAACACATTCATG	Klonierung
PB051380 KO 5' fwd	ATGCGGCCGATATGGTAATATATATGCTCATATTT TTCACAC	Klonierung
PB051380 KO 5' rev	ATGCAGCTGGACAGTATTCTTATTCTCAGTCAAAG G	Klonierung
PB082510 KO 3' fwd	ATGCTCGAGTGTCTATACAACAAAAAATTATGTAC CTTG	Klonierung
PB082510 KO 3' rev	ATGAAGCTTAAGAGGGTAAACATATAATAAATATT AACCC	Klonierung
PB082510 KO 5' fwd	ATGCGGCCGGAATTTTTTTTTTTTGTCCATCATT TCC	Klonierung
PB082510 KO 5' rev	ATGCAGCTGAATTAACATATGGGTAAATGAAACA ACG	Klonierung
PB112060 KO 3' fwd	GAGCTCGAGATTGAATATACCAAACAATAACAC CATC	Klonierung
PB112060 KO 3' rev	GAGATCGATCTCAATTTTAAATTACGTGCAAACG	Klonierung
PB112060 KO 5' fwd	GAGCGGCCGTTTTACGGCGAGTTGTAAGC	Klonierung
PB112060 KO 5' rev	GAGGAATTCGGCTATTCTTCAATACACCTGATG	Klonierung
PB121000 KO 3' fwd	ATGCTCGAGGATTTGCATATTTTATCCGCTGTG	Klonierung
PB121000 KO 3' rev	ATGATCGATATTTAAGTTATTGCCAATTGATAAATA AGACAC	Klonierung
PB121000 KO 5' fwd	ATGCGGCCGTGGAAGAATATTCGATCATACAAGG	Klonierung
PB121000 KO 5' rev	ATGCAGCTGATGGAAATTATGTATGTCCCCAC	Klonierung
PB145840 KO 3' fwd	GAGCTCGAGGCATGTTTATATCAATATTGTAATG GCTC	Klonierung
PB145840 KO 3' rev	GAGATCGATGTGGCAAAGCTACGAAATGATAG	Klonierung
PB145840 KO 5' fwd	GAGCGGCCGGTCTACATTAAGGGTCATTAGTT G	Klonierung
PB145840 KO 5' rev	GAGCAGCTGATAATTTGTCAACATTATTTAAACA AAAAATG	Klonierung
<b>Genotypisierung für Deletionskonstrukte</b>		

# Anhang

PB0198 3 Geno PB407	ATGAAATACCGCTCCATTTTTC	Kolonie-PCR
PB080730 3Geno In	GCATCATAACAATATACCATTTTACGG	Genotypisierung
PB080730 3Geno Out	CCTCCACATAATATACATCATCTATAGC	Genotypisierung
PB080730 5Geno In	CATACACAACATCGTCATGTACC	Genotypisierung
PB080730 5Geno Out	GTAGTTTTTATGATAGTTGGATTTCCTTG	Genotypisierung
PB143670 3Geno In	ATCGGCTGAACAACTGAAAC	Genotypisierung
PB143670 3Geno Out	GCTCAAACATCTAGATTTAAAGGTG	Genotypisierung
PB143670 5Geno In	CCCTTGTTAACAATACATTTGAGGG	Genotypisierung
PB143670 5Geno Out	CCTATTTGAGAATCCTTTCCACC	Genotypisierung
PB051680 3Geno In	CCTAAAAGGATAAAAAATCAGCCAATAC	Genotypisierung
PB051680 3Geno Out	CCCTCGAAATAAATTTCCCTTTTCC	Genotypisierung
PB051680 5Geno In	CCCCTTCCAAGCAATATCCC	Genotypisierung
PB051680 5Geno Out	GGTTCCTCATTATGATGATAAAGTCCATC	Genotypisierung
PB020810 3Geno In	ATGCAGATCAGAAAAAGAACAGC	Genotypisierung
PB020810 3Geno Out	CACATTTTCCAAGTGTAATTTGAAGG	Genotypisierung
PB020810 5Geno In	CACATTTCCGACTTCACATTTG	Genotypisierung
PB020810 5Geno Out	GCGGCTTGTTTATAATTTTATTTTCCAG	Genotypisierung
PB142250 3Geno In	CCCGAACCATGAAAAATAATTTTATGG	Genotypisierung
PB142250 3Geno Out	GTGTGGTGGGAAAAAACATG	Genotypisierung
PB142250 5Geno In	GTAATGTTAATAGCACAATTGCTGATAC	Genotypisierung
PB142250 5Geno Out	GGTAATATCACCAGTTTCAGATTTTCG	Genotypisierung
PB092630 3Geno In	CGATAGAATTTTAGGAGTTCTTGATGC	Genotypisierung
PB092630 3Geno Out	AGTATTTTCCATTGGGATTAACCTC	Genotypisierung
PB092630 5Geno In	AACTTGCAATATGTATCATGCTTCC	Genotypisierung
PB092630 5Geno Out	CCCATGGTGCAATAAATTTGTTTTATG	Genotypisierung
PB124410 3Geno In	GCAAATGAATGTATATAACGAAGATATCTATG	Genotypisierung
PB124410 3Geno Out	GACGTAGCCTAATGGATAAGGC	Genotypisierung
PB124410 5Geno In	TCCAAATCGTCTTTTTACTGTGC	Genotypisierung
PB124410 5Geno Out	CACACATGTTAAGGGACTTTACTATG	Genotypisierung
PB081630 3Geno In	TGTATGCAGATTGGATTTGCTCC	Genotypisierung
PB081630 3Geno Out	CATCTTTAAAGGATGAATGTGCAG	Genotypisierung
PB081630 5Geno In	GGATACCTGCCAAGGTATCATAAG	Genotypisierung
PB081630 5Geno Out	CCGAATTTGATTGTAAATTAGCATATGC	Genotypisierung
PB144130 3Geno In	GAAAAATTGTATCTAAATTTTGTGCC	Genotypisierung
PB144130 3Geno Out	CTCTTTCTCCAAATATCCAGTGTC	Genotypisierung
PB144130 5Geno In	CAATATCTTATTCTATATCTATGCCAAAGGC	Genotypisierung
PB144130 5Geno Out	GTTAGCTGAAAAACAAATATGTTACAC	Genotypisierung
PB040990 3Geno In	CTAAGCAAGGATTTTCATCAATACCAG	Genotypisierung
PB040990 3Geno Out	CAACTACTCAATGACCCCTTAC	Genotypisierung
PB040990 5Geno In	TGTGTTTCAAATTTGAGTCGAGC	Genotypisierung
PB040990 5Geno Out	GTAATATATGTGAAATATATTTTGCAATTCCTAAC	Genotypisierung
PB051380 geno 3' g fwd	CCATTAGGAAAAAATTTAATACACACTCC	Genotypisierung
PB051380 geno 3' rev	CTATTCTCATTTAGGTTGTCTATTTTGC	Genotypisierung
PB051380 geno 5' fwd	CTTCAAACGTTAATCTTCTCATTTTTC	Genotypisierung
PB051380 geno 5' g rev	GCATACGTTTCGCTTTATATTTTATTCTAC	Genotypisierung
PB082510 geno 3' g fwd	TATATGATGATGATACAAATTTTGAGCAC	Genotypisierung



## Anhang

PB082510 geno 3' rev	ATATTTATATTAGGAGCTACAGGAAATCG	Genotypisierung
PB082510 geno 5' fwd	CAATGAATATATTTCCACCTTTCATGC	Genotypisierung
PB082510 geno 5' g rev	GTCTATAAACTAACATCATTGGTTGC	Genotypisierung
PB112060 geno 3' g fwd	ATTTGCCAAATAATGTAGAACAAAAGG	Genotypisierung
PB112060 geno 3' rev	AGTATTTAGTACATAATTTAGGATCTGGC	Genotypisierung
PB112060 geno 5' fwd	TGTTTTAAATTTGTTCAAGTTTTAGTGTG	Genotypisierung
PB112060 geno 5' g rev	TATCACTAATTAGTGAGTTGATATACGC	Genotypisierung
PB121000 geno 3' g fwd	ATAAATAATTATGATACCTGCTCAATTCAG	Genotypisierung
PB121000 geno 3' rev	TAACATAATAAGCATGCTTGAAATAACG	Genotypisierung
PB121000 geno 5' fwd	CACTTATATTCGTGATATGTAATATATGGG	Genotypisierung
PB121000 geno 5' g rev	GGAATTCATATGCTTAATCTTAATCTCG	Genotypisierung
PB145840 geno 3' g fwd	AATACAAATGATGTTAACAGAACTTACC	Genotypisierung
PB145840 geno 3' rev	GCTAATTCATCTTATTCAATCTCCAACC	Genotypisierung
PB145840 geno 5' fwd	CAATCATGCAATAATCATATAGTATGTGTG	Genotypisierung
PB145840 geno 5' g rev	AGCACACTCTTTCTTAACTCTATAAAC	Genotypisierung

## Verwendete Software

Anhang Tabelle 3: Liste der verwendeten Programme und Software. Für frei erhältliche Programme ist die entsprechende Publikation angegeben.

Software	Version	Hersteller
SnapGene	2.8.2	GSL Biotech LLC
ImageLab	5.2.1	Bio-Rad
Microsoft Office	2016	Microsoft
QuantityOne	4.6.3	Bio-Rad
Adobe Illustrator	21.0.2	Adobe
CLC Workbench	8.0.1	Qiagen
ImageJ Fiji	1.52i	Schindelin et al. 2012
Integrated Genome Browser	9.0.1	Freese et al. 2016
7500 Software	2.3	Applied Biosystems
EndNote	X8.2	Clarivate Analytics
BLAST	-	NCBI
Mfold	-	Walter et al. 1994, Mathews et al. 1999
TargetP	1.1	Emanuelsson et al. 2000
MitoProt II	-	Claros und Vincens 1996
PlasmoAP	-	Foth et al. 2003
PATS	-	Zuegge et al. 2001

### Liste der PPR-Protein

Anhang Tabelle 4: Liste der PPR-Proteine für die Erstellung eines Suchmotivs.  
Registernummern der Proteine für <https://www.uniprot.org/>.

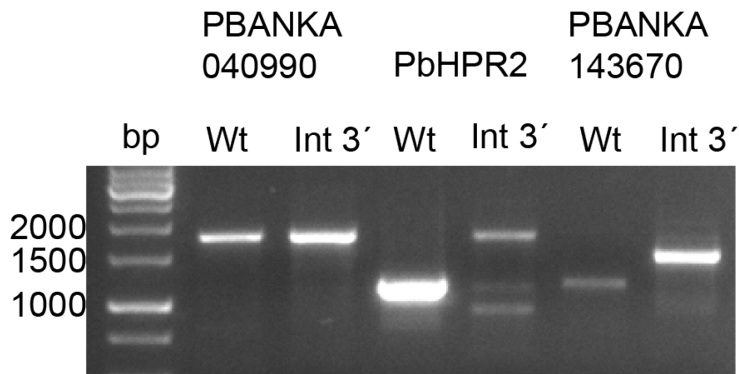
Proteinname	Organismus (Uniprot Nummer)
At1g64583	A. thaliana (P0C7R3)
At1g69290	A. thaliana (P0C7R4)
MRL1	A. thaliana (Q0WLC6)
At1g66345	A. thaliana (Q3ECH5)
HCF152	A. thaliana (Q9SF38)
At1g63400	A. thaliana (Q9SH26)
PPR566-6	Z. mays (B6SYZ1)
PPR10	Z. mays (B8Y6I0)
PPR5	Z.mays (A7LN87)
PPR-814b	Z.mays (C9W4C1)

## Sequenzen der cosRNAs

Anhang Tabelle 5: Sequenzen der identifizierten cosRNAs.

Name der cosRNA	Genomposition (erstes und letztes nt)	Länge [nt]	Genom-strang	Sequenz 5'-3'
cosRNA1	389 - 355	35	-	TTTGAACCTTGAACAAGGTTCCATTGGAATGAGAGT
cosRNA2	389 - 333	57	-	TTTGAACCTTGAACAAGGTTCCATTGGAATGAGAGTTCA CCGTTAGAAGCGATACGTG
cosRNA3	1515 - 1554	40	+	TAAACATATAACGGTAAGAAGGTTCCCGGGGATAAC AGG
cosRNA4	1698 - 1725	28	+	GAACTCCAGGCGTTAACCTGTAGAGTTG
cosRNA5	2030 - 1972	59	-	TGAGGCTCGGATATAAATGAGAAATATTTGATCCATAC AGTCCCAGCGACAGCGGTTAT
cosRNA6	4849 - 4868	21	+	GGAGTTGGCAAGTTAAAAAAA
cosRNA7	4893 - 4944	52	+	ATTAATGTTAGGATGTATGGGATATTTGTAGTACACCTT GATTGGTTTTACT
cosRNA8	5201 - 5157	45	-	TTATAGCCATGTCTCCATGAACTATAAAACATGTGATCT AATTAC
cosRNA9	5283 - 5243	41	-	TTGAAGTGTGGAGAGAATCCTCTTAGTAACCAACATC TGG
cosRNA10	5283 - 5226	58	-	TTGAAGTGTGGAGAGAATCCTCTTAGTAACCAACATC TGGAATCGAGAGAGATTCC
cosRNA11	5507 - 5478	30	-	AAACAGTCGGTGCGAAGTCGTAACATGGTA
cosRNA12	5771 - 5746	26	-	TGATAAACGGCGGCTGTATTTTAAAC
cosRNA13	166 - 205	40	+	CATAAGAATAGAAACAGATGCCAGGCCAATAACTCAA AA
cosRNA14	787 - 725	63	-	CATTTTGTGAAATAGTCTGGTTATTTATAGAGTTCTTA TTCTATTCAAGATAAATGAACATG
cosRNA15	1492 - 1513	22	+	AAGAACGGTGAGATAATGTGCC
cosRNA16	1630 - 1601	30	-	TTTATTGCAAGGCTGCGATGAGACGACATG
cosRNA17	1638 - 1725	88	+	TAGCGTGATTGTTGCCTTGTACACACCGCTCGTCACG CAATATCAATATACTGGGTATAGAACTCCAGGCGTTAA CCTGTAGAGTTG
cosRNA18	2067 - 2031	37	-	AATAAGTGAATATCTATTTAAACAATAAAATAACA
cosRNA19	3480 - 3537	58	+	TTTATTGTTTTATGAACCTTTACTCTATTAATTTAGTTA AAGCACACTTAATAAATT
cosRNA20	4866 - 4944	79	+	GAAGTTCTGGTTTATAATAGATACGTTATTAATGTTAGG ATGTATGGGATATTTGTAGTACACCTTGATTGGTTTTAC T

## Überprüfung der korrekten Integration der Markierungskonstrukte ins Genom von *P. berghei*



Anhang Abbildung 1: PCR Ergebnisse der Überprüfung der Genomintegration von Markierungskonstrukten. Die PCR wurde mit genomischer DNA als Vorlage durchgeführt. Alle Oligonukleotide wurden erstellt um die Integrationsregion zu amplifizieren. Die unveränderte Genomsituation (Spur Wt) ist in allen drei Konstrukten erkennbar und entspricht den erwarteten Größen. Die Integration wurde mittels eines genspezifischen Oligonukleotids und einem vektorspezifischen Oligonukleotid überprüft (Int 3'). Auch hier wurden Banden in der entsprechenden Größe detektiert. Ein Schema der Oligonukleotide ist in Abbildung 16 zu finden.

## Vorhergesagte helikale Elemente von PbHPR1 und PbHPR2 in Relation zu den identifizierten HPR-Motiven

Anhang Abbildung 2: Die Strukturvorhersage von PbHPR1 und 2 in Relation zu den identifizierten HPR-Motiven. Die Primärsequenz (SEQ) sowie die Sekundärstrukturelemente (SS) der Proteine ist dargestellt. In der SS sind Aminosäuren, die in der Vorhersage Teil eines helikalen Elementes sein könnten, mit H markiert. Die identifizierten HPR-Motive sind abwechselnd doppelt und einfach unterstrichen. Der rASA-Wert (*relative accessible surface area*) ist ein Maß für die potentielle Zugänglichkeit der AS an der Proteinoberfläche. Verdeckte AS (rASA <20%) sind in blau dargestellt.

### PBANKA\_051680 = PbHPR1

```

SEQ: 1  MLIYGILLGRGKANKIYPQNIGRWMYAHFVTCTTNDNIILKRNYEEIKNL 50
SS : 1  --H-HHH-----HHHHHHHHHHHH-----HHHHHHHHHHHH-- 50
rASA: 1  55444444556528714131002001011232544353015303620671 50

SEQ: 51  KAQELRRLSESCCIKKISDVIIWSEICRNAIKKSDEFKHFDAALLLSCFD 100
SS : 51  -HHHHHHHHHHHHHH-----HHHHHHHHHHHHHH-H-H--HHHHHHHHHHHH 100
rASA: 51  61420230021003371634500330052026216505241012001002 100

SEQ: 101 KMNLLDKILYTHFSDIFIKNNINFEPRLILLINLYCKVNIFFPRILFIEV 150
SS : 101 H-----HHHHHHHHHHHHHH-----HHHHHHHHHHHHHH-----HHHHHH 150
rASA: 101 52614454003200510272076041410010010014161335400230 150

SEQ: 151 FHAIIRYSSKLYPNYVDIFECFAKYEIANKDLISTLCKSIVKNINLFGY 200
SS : 151 HHHHHH-H---HHHHHHHHHHHHHH-----HHHHHHHHHHHHHH-H-H--H 200
rASA: 151 03101510450313000100100042614356004300610262074032 200

SEQ: 201 TDLCSIVGALRSLKINDDVFFYCIDKQKELKLMSTCQELFDYINKIKLL 250
SS : 201 HHHHHHHHHHHHHHH-----HHHHHHHHHHHHHHHH-----HHHHHHHHHHHHHH- 250
rASA: 201 41001001003418153530031015202610651424202300230331 250

SEQ: 251 KYSWELYEKDLIKEFLNRINDFKNGNDVNQLHDPFICLNLYISKNIISNN 300
SS : 251 ----HHHHHHHHHHHHHHHHHH-----HH-HHHHHHHHHHHHHHHHH-----HH 300
rASA: 251 52314301530151024105525556215302300100110153711354 300

SEQ: 301 FLLTSLMWCANQVYQYPSRSKRPLSYQLIKLYQIMKEHNVENYDFIEKA 350
SS : 301 HHHHHHHHHHHHHHH-----HHHHHHHHHHHHHHHH-----HHHHHHHH 350
rASA: 301 00310131014104524344254322420240141056361743520351 350

SEQ: 351 IHKFVISRGGLATNREKITKPVSYQKGRKYIFTPDPLNAYSED TNKYIDS 400
SS : 351 HHHHHH-----EEE-----EE----- 400
rASA: 351 02410354323344555345435246745111154425444554554346 400

SEQ: 401 SNKTENIPNKNEQLAYSNNMYQGNDKDFDLTETYSYNENANDHIKYNPNIL 450
SS : 401 -----HHHHHH----- 450
rASA: 401 7575664456655443454456666263454354565455436564543 450

SEQ: 451 TQKQKLINLSLEKKTENKKETYSNSRHCNFKLRQRPKRIKNQPIQDQK 498
SS : 451 -H-H----- 498
rASA: 451 455455263647765666766456554242635654753665657778 498

```

**PBANKA\_051380 = PbHPR2**

```

SEQ: 1  MGLLRPRERVIVNAIRRESRIKYKAKRMHKRFRSWAQQRVREYWLPMNVC 50
SS : 1  -----HHHHHHHHHHHHHH-H--HHHHHHHHHHHHHHHHHHHH---HHHH 50
rASA: 1  77756355632342245235454535544553543244315543332441 50

SEQ: 51  LTSQVNLMDGQYISACVQKAATLRKHDYALWKGYSERILEICNTLTPOOI 100
SS : 51  HH--HH---HHHHHHHHHHHHHH-----HHHHHHHHHHHHHH-----HHHH 100
rASA: 51  24342640415201200330142735445004200520251075032410 100

SEQ: 101 GYIFYGMGKSGFLNMPFYNI FLNNVENCLNFYSHPLMCVAWALNRLQIR 150
SS : 101  HHHHHHHHHH-----HHHHHHHHHHHHHH-H-----HHHHHHHHHHHHHH----- 150
rASA: 101  020011005333133350042005103610770413100100100142616 150

SEQ: 151 KEEFLKKLCKCVINKFDEIRIKDLIKINTSIAKLGIDDKNYKTFINKNII 200
SS : 151  -HHHHHHHHHHHHHH-H-----HHHHHHHHHHHHHH-----HHHHHHHHHHHH 200
rASA: 151  45600430051016206504151013001101415163640132015201 200

SEQ: 201 NKLETIFAQDFRNVINDVTLINLYDDEVKKYILTRFSNMFICARPOHYKS 250
SS : 201  HHH-----HHHHHHHHHHHHHHHHHH--HHHHHHHHHHHH-HHHHH--HHHHH 250
rASA: 201  520550313101100000000011113301320132023111213132021 250

SEQ: 251 AYKSAVSIRVLYPHVWDSLTKVKSFYVRLSMRRIPEIPRTPSAFQWEVS 300
SS : 251  HHHHHHHHHH--HHHH-H--HHHHHHHHHH-----HHHHHHH 300
rASA: 251  00200010111045005312650331124124551452343323312301 300

SEQ: 301 DCLAKLGIGHRNTFLWGCYFIDIGETNEKRNCWFVDGPSSFYTATNQYTE 350
SS : 301  HHHHH--EEEEEEEE--EEEE-----EEEE-----H 350
rASA: 301  31046250214220212000010133553300000112311234364223 350

SEQ: 351 SVKLQHRILYDLGWNIRRIWIDWLNMGNDICQKIEFIKKRNNEPLGKN 400
SS : 351  HHHHHHHHHHH--EEEE-HHHHHH---HHHHHHHHHHHH----- 400
rASA: 351  21411230055271412502131056247535411410352276654365 400

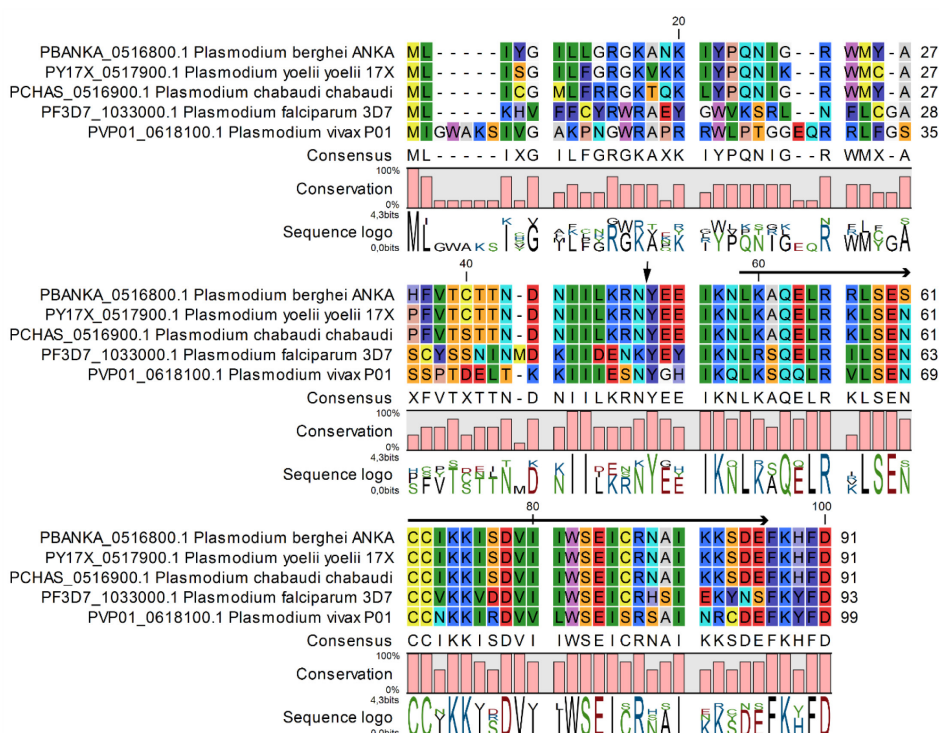
SEQ: 401 LIHTPVIHPDEIASKLKLKLYKSNKEMEKEYKNKERIELNM 441
SS : 401  -----HHHHHHHHHHHHHHHHHHHHHHHHHH----- 441
rASA: 401  24556615474135326526555566666656767646466 441

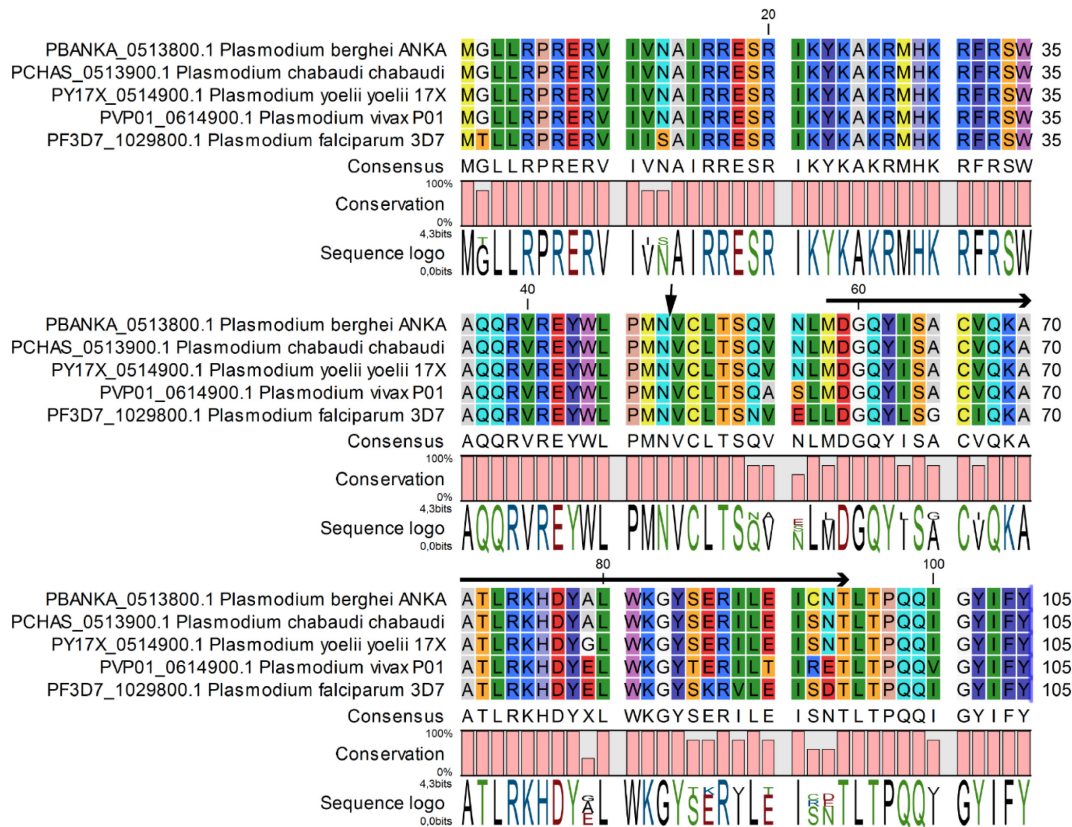
```

## Identifikation des potentiellen mitochondrialen Lokalisationspeptids in PbHPR1 und PbHPR2

Anhang Abbildung 3: Um das mitochondriale Lokalisationspeptid zu identifizieren, wurde die N-terminale Sequenz (91 AS) des PbHPR1 bzw. 105 AS des PbHPR2 aus *P. berghei* mit den Orthologen anderer Plasmodium Spezies (*P. yoelii*, *P. chabaudi*, *P. falciparum* und *P. vivax*) verglichen. Die Startposition des rekombinanten Proteins ist mit einem schwarzen Dreieck gekennzeichnet. Die Position des ersten HPR-Motivs ist durch einen Pfeil gekennzeichnet.

### PbHPR1:



**PbHPR2:**



## Abkürzungen

AS	Aminosäure
°C	Grad Celsius
A, T, C, G, U	Basen der DNA bzw. RNA: Adenin, Cytosin, Guanin, Thymin, Uracil
<i>A. dest</i> / <i>A. bidest</i>	deionisiertes Wasser ( <i>Aqua destillata</i> , <i>Aqua bidestillata</i> )
APS	Ammoniumperoxodisulfat
AS	Aminosäure
ATP	Adenosintriphosphat
bp	Basenpaare
BSA	Rinderserum Albumin
bspw	Beispielsweise
bzw	beziehungsweise
<i>C. reinhardtii</i>	<i>Chlamydomonas reinhardtii</i>
ca.	zirka
cDNA	<i>complementary</i> DNA
CDS	kodierende Sequenz
Ci	Curie
cosRNA	clustered organellar short RNA
dATP	Desoxyadenosintriphosphat
dCTP	Desoxycytidintriphosphat
dGTP	Desoxyguanosintriphosphat
DNA	Desoxyribonukleinsäure
dNTP	gemisch aus dATP, dTTP, dGTP und dCTP
DTT	Dithiothreitol
dTTP	Desoxythymidintriphosphat
<i>E. coli</i>	<i>Escherichia coli</i>
EDC	1-Ethyl-3-[3-dimethylaminopropyl]-carbodiimid
EDTA	Ethylendiamintetraessigsäure
<i>et al.</i>	<i>et alii</i> (unter Anderen)
EtBr	Ethidiumbromid
EtOH	Ethanol
FASTK	<i>Fas-activated serine/threonine kinase</i>
g	Erdbeschleunigung bzw. Gramm
GFP	green fluorescent protein
GVO	gentechnisch veränderter Organismus
HPR	Heptatricopeptide <i>repeat</i>
HRP	horseradish peroxidase (Meerrettichperoxidase)
HTS	Hochdurchsatzsequenzierung
IgG	Immunglobulin G
inkl.	inklusive
IP	Immunpräzipitation
J	Joule
kb	Kilobasenpaare
l	Liter
M	Mol (mol/l)
MBP	Maltosebindeprotein
min	Minuten

## Abkürzungen

MOPS	3-(N-Morpholino)propansulfonsäure
mRNA	messenger RNA
Na	Natrium
NaCl	Natriumchlorid
NaOH	Natriumhydroxid
NEP	nukleär kodierte plastidäre RNA-Polymerasen
nt	Nukleotid
OPR	octatricopeptide <i>repeat</i>
ORF	offener Leserahmen
P. berghei	Plasmodium berghei
P. falciparum	Plasmodium falciparum
PAGE	Polyacrylamid-Gelelektrophorese
PCR	Polymerase-Kettenreaktion
PEP	plastidär kodierte RNA-Polymerase
PPR	pentatricopeptide <i>repeat</i>
psi	<i>pounds per square inch</i>
RNA	Ribonukleinsäure
rpm	Umdrehung pro Minute
rRNA	ribosomale RNA
RT	Raumtemperatur
s	Sekunden
SDS	Natriumdodecylsulfat
SSC	Natriumchlorid-Natriumcitrat-Puffer
TAE	Tris-Acetat-EDTA-Puffer
TBE	Tris-Borat-EDTA-Puffer
TBST	Tris gepufferte Saline Tween 20
TCA	Tricarbonsäurezyklus
TEMED	Tetramethylethylendiamin
TEV	<i>Tobacco Etch Virus</i>
Tris	Tris-[hydroxymethyl]-aminomethan
tRNA	transfer RNA
UTR	nicht translatierter Bereich
UV	Ultra-violetes Licht
V	Volt
v/v	Volumen pro Volumen
Vol.	Volumen
w/v	Gewicht pro Volumen
WT	Wildtyp
z.B.	zum Beispiel

## Danksagung

Meine ersten Dankesworte gehen an meinen Chef Prof. Dr. Christian Schmitz-Linneweber, für die ausgezeichnete Betreuung und die konstante Unterstützung. Nachdem im ersten Jahr meiner Promotion mein Pflanzenprojekt wenig Erfolg hatte, hat er mir die Chance gegeben mit dem Malaria-Projekt ein neues Feld im Labor zu eröffnen. Ein solcher Vertrauensvorschuss ist nicht selbstverständlich und ich bin sehr dankbar dafür.

Ich danke auch Prof. Dr. Kai Matuschewski, nicht nur für die Infrastruktur die er bereitgestellt hat und die Begutachtung meiner Arbeit, sondern auch für seine Hilfsbereitschaft und sein Interesse an dem Projekt. Erst durch seine Unterstützung waren die Arbeiten mit Malaria möglich.

Mein Dank gilt auch Dr. Elena A. Levashina und ihren Mitarbeitern, die mich gerade am Anfang mit *Plasmodium falciparum* versorgt haben und dadurch die Arbeiten mit dem humanen Erreger ermöglicht haben.

Ich danke Prof. Dr. Volker Knoop für die Begutachtung meiner Arbeit.

Ich danke Joachim und Katja von der AG Matuschewski, die mir bei den Arbeiten mit *Plasmodium berghei* und allen Mausexperimenten so großartig geholfen haben. Auch allen weiteren Mitarbeitern der Gruppe möchte ich danken, die mich so nett im Labor aufgenommen haben.

Meinen Doktorandenkollegen Marie, Stephie, Gongwei, Ben und Birte möchte ich herzlich danken für die schöne Zeit, die Diskussionen und die Toleranz für eine ansehnliche Anzahl fragwürdiger Äußerungen meinerseits. So eine Doktorandenzeit wäre alleine in der Tat langweilig. Hannes möchte ich danken für seine Geduld und Unterstützung mit Ideen und Rat, obwohl ich nicht mehr sein Masterstudent bin. Julia, Nadine, Jose, Prof. Börner und allen weiteren Leuten der Molekularen Genetik danke ich für die Diskussionen nebenbei und die ganzen Kleinigkeiten die ein Labor angenehm machen. Das gleiche gilt für alle Studenten und ehemaligen, deren Namen ich hier nicht alle aufführen kann. Unseren

Assistenten Conny, Irina, Shumon, Carola, Tabea und Jana danke ich für die Organisation und Hilfe in allen technischen Dingen.

Ich danke meinen Studenten Martin, Sabrina und Franziska die ich im Labor betreuen durfte. Martin hat mich bei der Lokalisation von Proteinen unterstützt und einen Array entworfen, der leider (noch) nicht zum Einsatz gekommen ist. Sabrina hat weiter untermauert, dass die HPRs doch sehr wichtig sind und gleichzeitig einen Gradienten entwickelt der hoffentlich die mitochondrialen Ribosomen unter die Lupe nehmen kann. Franziska arbeitet an der Deletion der verbleibenden HPRs und versucht gleichzeitig weitere HPRs in *E. coli* zu exprimieren.

Außerhalb der Labore danke ich meiner Familie und meinen Freunden für die Geduld und Unterstützung. Durch sie wurde eine Menge Frustration abgefangen. Meiner Freundin Sabine danke ich für die konstante Aufmunterung und Ablenkung, die engelsgleiche Geduld mit mir und den Arbeitszeiten. Es ist auch nicht immer leicht mit so einem Wissenschaftler und beinhaltet viele verschobene Termine und Verabredungen.

Allen denen die ich hier nicht namentlich erwähnt habe, sei ein allgemeines „Danke“ ausgesprochen mit der Bitte es ganz persönlich auf sich beziehen.

## **Selbstständigkeitserklärung**

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben. Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Doktorgrad. Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde. Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsordnung der Lebenswissenschaftlichen Fakultät (ehemals Mathematisch-Naturwissenschaftliche Fakultät I) der Humboldt-Universität zu Berlin vom 27. Juni 2012

Berlin, 22.03.2019

---

Arne Hillebrand